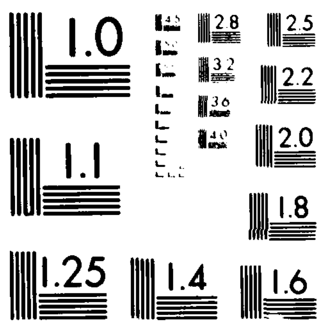
AD-A083 808 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
MOVING-WEIGHTED-AVERAGE SMOOTHING EXTENDED TO THE EXTREMITIES O--ETC(U)
NOV 79 T N GREVILLE DAAG29-75-C-0024
UNCLASSIFIED MRC-TSR-2025 NL

END
DATE FILMED
6-80
DTIC


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

LEVEL II (2) B.S.

ADA083808

9 MRC Technical Summary Report, # 2025

6 MOVING-WEIGHTED-AVERAGE SMOOTHING EXTENDED TO THE EXTREMITIES OF THE DATA.

10 T. N. E./Greville

Revision.

14 MRC-TSR-2025

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

11 Nov[redacted]79 12 69

Received October 10, 1979

15 DAAG29-75-C-0024

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709

80 4 9 108

221200

ERRATA

MRC Technical Summary Report #2025

MOVING-WEIGHTED-AVERAGE SMOOTHING
EXTENDED TO THE EXTREMITIES OF THE DATA

T. N. E. Greville

Location	Actual reading	Correct reading
Page 19, bottom line, upper limit of summation	$m+s$	$m-s$
Page 24, equation (8.10)	$(-1)^s$	$g(-1)^s$
Page 24, last displayed equation	$Q - j$	$Q - j + 1$
Page 55, Table 6:		
$i = 1,\ j = 1$	-.000046	.001040
$i = 1,\ j = 4$	-.064956	-.089936
$i = 2,\ j = 1$	.130190	-.073930

On page 13 it would be desirable to replace the line following equation (5.3) by the following:

converges in some part of the complex plane. (Note that it follows from properties (1)-(4) that t_ν depends only on ν and is independent of N.)

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

MOVING-WEIGHTED-AVERAGE SMOOTHING EXTENDED TO THE EXTREMITIES OF THE DATA[†]

T. N. E. Greville

Dedicated to the memory of Erastus Lyman De Forest (1834-1888)

Technical Summary Report #2025
November 1979

ABSTRACT

The use of a symmetrical moving weighted average of $2m + 1$ terms to smooth equally spaced observations of a function of one variable does not yield smoothed values of the first m and the last m observations, unless additional data beyond the range of the original observations are available. By means of analogies to the Whittaker smoothing process and some related mathematical concepts, a natural method is developed for extending the smoothing to the extremities of the data as a single overall matrix-vector operation having a well defined structure, rather than as something extra grafted on at the ends.

Key Words: Smoothing; Moving weighted averages; Extrapolation; Toeplitz matrices; Graduation.

AMS(MOS) Subject Classification: 65D10, 62M10, 62P05, 65F30

Work Unit Number 3: Numerical Analysis and Computer Science

[†]Revision of Technical Summary Report #1786. - A046380

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024

SIGNIFICANCE AND EXPLANATION

The use of a moving weighted average of $2m + 1$ terms to smooth equally spaced observations of a function of one variable does not yield smoothed values of the first m and the last m observations, unless additional data beyond the range of the original observations are available. Using Toeplitz matrices, Laurent series, and analogies to the Whittaker smoothing process, we develop a natural method of extending the smoothing to the extremities of the data.

MOVING-WEIGHTED-AVERAGE SMOOTHING
EXTENDED TO THE EXTREMITIES OF THE DATA

T. N. E. Greville

1. INTRODUCTION

A time-honored method of smoothing equally spaced observations of a function of one variable to remove or reduce unwanted irregularities is the moving weighted average (MWA). An example is Spencer's 15-term average (Macaulay 1931; Henderson 1938), which can be expressed in the form

$$u_x = \frac{1}{320}(-3y_{x-7} - 6y_{x-6} - 5y_{x-5} + 3y_{x-4} + 21y_{x-3} + 46y_{x-2} + 67y_{x-1}$$
$$+ 74y_x + 67y_{x+1} + 46y_{x+2} + 21y_{x+3} + 3y_{x+4} - 5y_{x+5} - 6y_{x+6} - 3y_{x+7}), \tag{1.1}$$

where y_x is the observed value corresponding to the argument x, and u_x is the corresponding adjusted value. Actuarial writers commonly refer to such smoothing as "graduation."

More generally (Schoenberg 1946) a symmetrical MWA is of the form

$$u_x = \sum_{j=-m}^{m} c_j\, y_{x-j}\ , \tag{1.2}$$

where m is a given positive integer and the real coefficients c_j are such that $c_{-j} = c_j$ and

$$\sum_{j=-m}^{m} c_j = 1\ .$$

Such averages have a long history that is not widely known. One of the earliest writers on the subject was the Italian astronomer G. V. Schiaparelli (1866), who is remembered chiefly for his observations of the planet Mars. Futher contributions were made by the Danish actuary and mathematician J. P. Gram and the Danish astronomer T. N. Thiele, both of whom played major roles in the early development of statistical theory. The majority of publications on this subject have appeared in English and Scottish actuarial journals starting with John Finlaison in 1829 (see Maclean 1913). Probably the first writer to make a systematic investigation of such averages was the American mathematician E. L. De Forest (1873, 1875, 1876, 1877). His work,

Revision of Technical Summary Report #1786.

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

published in obscure places, was rescued from total oblivion largely through the efforts of Hugh H. Wolfenden (1892-1968), who also made important contributions to the subject (Wolfenden 1925). E. T. Whittaker (1923) suggested an alternative method of smoothing, which has been widely employed, especially by actuaries, and will be referred to extensively later, because of numerous analogies to the MWA procedure. The first writer to apply sophisticated mathematical tools to the study of these averages was I. J. Schoenberg (1946, 1948, 1953), who introduced the notion of the characteristic function of an MWA, and utilized it to formulate a criterion for judging whether a given average can properly be called a "smoothing formula." This criterion will be discussed in Section 11.

2. THE PROBLEM OF SMOOTHING NEAR THE EXTREMITIES OF THE DATA

When MWA's have been used by actuaries, the argument x is usually age (of a person) in completed years. When they are used for smoothing economic time series, x denotes the position of a particular observation in a time sequence. The latter area of application appears to stem largely from the work of Frederick R. Macaulay (1931), who was the son of an actuary.

In either case, a serious disadvantage of the method is that it does not produce adjusted values for arguments too near the extremities of the data. For example, suppose Spencer's 15-term average is used to smooth monthly data extending from 1970 through 1976. The formula does not give smoothed values for the first 7 months of 1970 or the last 7 months of 1976 unless data can be obtained for the last 7 months of 1969 and the first 7 months of 1977. Clearly, acquisition of data extending farther into the past is less of a problem than acquisition of future data.

Actuaries in North America seem to have largely abandoned the use of MWA's in favor of Whittaker's method, which does not have the disadvantage

described. It is likely that British actuaries may still use these averages to some extent. They appear to be currently employed by economic and demographic statisticians (Shiskin and Eisenpress 1957; Shiskin, Young, and Musgrave 1967).

Various suggestions have been made (De Forest 1877; Miller 1946; Greville 1957, 1974a) for dealing with the problem of adjustment of data near the extremities, but none of them has won general acceptance. De Forest's (1877, p. 110) suggestion is so relevant to the subject of the present paper that it is worth quoting in full:

"As the first m and the last m terms of the series cannot be reached directly by the formula [of $2m + 1$ terms], the series should be graphically extended by m terms at both ends, first plotting the observations on paper as ordinates, and then extending the curve along what seems to be its probable course, and measuring the ordinates of the extended portions. It is not necessary that this extension should coincide with what would be the true course of the curve in those parts. The important part is that the m terms thus added, taken together with the $m + 1$ adjacent given terms, should follow a curve whose form is approximately algebraic and of a degree not higher than the third."

Elsewhere (Greville 1974a) I have proposed extrapolating the observed data by fitting a least-squares cubic to the first $m + 1$ values and a similar cubic to the last $m + 1$ observations. Though my proposal was made before I had noted the passage just quoted from De Forest, it is very much in the spirit of his suggestion; it is not a long step from graphic to algebraic extrapolation.

Another approach (Greville 1957) regards the adjustment process as a matrix-vector operation. We write

$$u = Gy , \tag{2.1}$$

where y is the vector of observed values, u is the corresponding vector of adjusted values, and G is a square matrix. If a specified symmetrical MWA of $2m + 1$ terms is to be used wherever possible, then the nonzero elements of G, except for the first m and the last m rows, are merely the weights in the moving average, these weights moving to the right as one proceeds down the rows of the matrix. In the first m and the last m rows special unsymmetrical weights, determined in some appropriate manner, must be inserted. The matrix approach and the extrapolation approach are not wholly unrelated, since the final results of the extrapolation approach can be expressed in matrix form.

It is the purpose of the present paper to show that when a given symmetrical MWA is being employed and fulfills certain minimal requirements, there is a natural, preferred method of extending the adjustment to the extremities of the data, strongly suggested by the mathematical properties of the weighted average. This natural method of extension seems to have eluded previous writers on the subject, as indeed it eluded me during the many years I have thought about the matter. The preferred method of extension has the interesting property that it can be arrived at either through the matrix approach or the extrapolation approach. In the latter case, one must employ a special extrapolation formula uniquely determined by the given MWA. Though the two approaches appear to be quite different, they will be shown in Section 8 to be mathematically equivalent, and they will give identical results except for rounding error. In the matrix approach the treatment of the values near the ends becomes an integral part of a single overall operation, and not something extra grafted on at the ends. It is especially fitting that it should be published now, since the centennial of De Forest's death occurs in the 1980's.

In my own thinking I arrived at the procedure first through the matrix approach, guided largely by extensive analogies to the Whittaker process (which is most conveniently expressed in matrix terms). It was only later that I became aware that identical results could be obtained by means of an extrapolation algorithm. Though the matrix approach provides greater insight into the rationale behind the procedure, the extrapolation approach is simpler computationally. Therefore, we shall first describe and illustrate the extrapolation algorithm, and shall then motivate and justify the procedure by means of the matrix approach. This investigation has led to some interesting mathematical developments (Greville and Trench 1979; Greville 1979; Greville 1980) that have been published elsewhere in a more general context and will be cited here as the need arises.

The extrapolation approach is merely a computational short cut, and nearly always the extended values obtained by its use are highly unrealistic if regarded as extrapolated values of the function under observation. (The reader will note that the quotation from De Forest earlier in this section contains a similar admonition.) This fact is irrelevant, but has seriously "turned off" some users. Hereafter I shall therefore avoid the use of the words "extrapolate" and "extrapolation," and shall speak of "extension," "extended values," and "intermediate values."

It is emphasized that the procedure to be described (or any other procedure for completing the graduation) is recommended for use only when additional data extending beyond the range of the original observations are not available.

Probably some readers will be primarily concerned with the application of the method to numerical data, and will have less interest in its mathematical development. Such readers will find the information they require in

the following Sections 3 and 4. On the other hand, readers who may wish to pursue the mathematical derivation first and leave computational details later may skip Sections 3 and 4 and pass at once to Section 5.

3. THE EXTENSION ALGORITHM

A weighted average of the form (1.2) will be called exact for the degree r if it has the property that, in case all the observed values y_{x-j} in (1.2) should happen to be the corresponding ordinates of some polynomial $P(x - j)$ of degree r or less, then

$$u_x = y_x = P(x) ,$$

but there is some polynomial of degree $r + 1$ for which this is not the case. In other words, an average that is exact for the degree r reproduces without change polynomials of degree r or less, but not in general those of higher degree. If the weights are symmetrical, r must be odd, and we may write $r = 2s - 1$. This implies that $r < 2m + 1$, and therefore $s \leq m$.

For a simple (unweighted) average, $r = 1$. For the overwhelming majority of MWA's used in practice, $r = 3$. The preference for cubics has a long history. De Forest (1873, p. 281) suggests that "a curve of the third degree, which admits a point of inflexion ... is ... better adapted than the common parabola to represent the form of a series whose second difference changes its sign."

We shall use the notation of the calculus of finite differences, wherein E is the "displacement operator" or "shift operator" defined by

$$Ef(x) = f(x + 1) ,$$

and δ is the "central difference" operator defined by

$$\delta f(x) = f(x + 1/2) - f(x - 1/2), \qquad (3.1)$$

so that

$$\delta^2 f(x) = f(x + 1) - 2f(x) + f(x - 1) .$$

If the weighted average (1.2) is exact for the degree $2s - 1$, it can be written in the form

$$u_x = [1 - (-1)^s \delta^{2s} q(E)]y_x , \tag{3.2}$$

where $q(E)$ is of the form

$$q(E) = \sum_{j=-m+s}^{m-s} q_j E^j \tag{3.3}$$

with $q_{-j} = q_j$. In a typical smoothing formula $q(E)$ has only positive coefficients, but this is not necessarily the case. If $q(z)$ is multiplied by z^{m-s} to eliminate negative exponents, the resulting polynomial is of degree $2m - 2s$. Because of the symmetry of the coefficients, it is a reciprocal polynomial. In other words, if ρ is a zero of the polynomial, it follows that ρ^{-1} is a zero. In general, we shall make the assumption that this polynomial has no zeros on the unit circle. The case in which it does have such zeros is mainly of theoretical interest and is briefly referred to in Section 7.

Let $p(z)$ denote the polynomial of degree $m - s$ with leading coefficient unity whose zeros are the $m - s$ zeros of $z^{m-s}q(z)$ located within the unit circle. In general, some or all of these zeros are complex, but they must occur in conjugate pairs, so that $p(z)$ has real coefficients. Now we define a polynomial $a(z)$ of degree m and its coefficients a_j by

$$a(z) = (z - 1)^s p(z) = z^m - \sum_{j=1}^{m} a_j z^{m-j} . \tag{3.4}$$

Suppose the given data consist of $N = Q - P + 1$ given values extending from $x = P$ to $x = Q$. We assume that $N \geq 2m + 1$, so that at least one smoothed value is obtained by direct application of the given MWA. Then we obtain m intermediate values to the left of $x = P$ by successive application of the recurrence

$$y_x = \sum_{j=1}^{m} a_j y_{x+j}.$$

Similarly, m intermediate values to the right of $x = Q$ will be obtained by

the analogous recurrence

$$y_x = \sum_{j=1}^{m} a_j\, y_{x-j}\ .$$

Finally, application of the symmetrical MWA of $2m + 1$ terms to the $N + 2m$ observed and intermediate values gives adjusted values u_x for $x = P$, $P + 1,\ldots,Q$.

For example, Spencer's 15-term formula (1.1) can be expressed in the form (3.2) with $s = 2$, where

$$q(E) = \frac{1}{320}\,(3E^{-5} + 18E^{-4} + 59E^{-3} + 137E^{-2} + 242E^{-1} + 318 + 242E + 137E^{2} + 59E^{3} + 18E^{4} + 3E^{5}).$$

Using a computer program to find the zeros of $z^5 q(z)$, constructing the polynomial $p(z)$, and finally applying the formula (3.4), we obtain for Spencer's 15-term formula

$$a(z) = z^7 - .961572z^6 - .372752z^5 - .015904z^4 + .123488z^3 + .125229z^2 + .075887z + .025624.$$

The coefficients are rounded to the nearest sixth decimal place, except that the final digits of the coefficients of z^3 and z^2 have been adjusted by one unit to make the sum of the coefficients exactly zero.

Note that in the trivial case $s = m$, $q(z)$ is a constant and $p(z)$ is unity. Thus the algorithm reduces to extrapolation of the observed data by sth differences (i.e., by fitting a polynomial of degree $s - 1$ to the first s observations and a similar polynomial to the last s observations).

As a numerical illustration, Spencer's 15-term average has been applied to some meteorological data. Table 1 and Figure A show the observed and graduated values of monthly precipitation in Madison, Wisconsin in the years 1967-71. No adjustment has been made for the unequal length of the months.

4. TABLES OF MOVING-AVERAGE AND EXTENSION COEFFICIENTS

Tables 2 and 3 show the coefficients in the MWA and the corresponding extension coefficients (that is, c_j and a_j) for 21 weighted averages that have appeared in the literature. Table 2 is devoted to the class of averages known to actuaries as minimum-R_3 formulas and to economic statisticians as "Henderson's ideal" formulas. They are discussed more fully in Section 8. The values in Table 2 are shown to six decimal places. In both instances, a few final digits have been adjusted by one unit to make the sum exactly unity. The moving-average coefficients are given to the nearest sixth decimal place except for the slight adjustments mentioned; rounding error in the computation of the extension coefficients may have introduced further small errors in some instances.

Table 3 is concerned with 11 moving averages derived by various writers on an ad hoc basis and known by the names of their originators. The source notes for this table do not attempt to cite the earliest publication of the formula in question, but merely indicate a convenient reference where it can be found. All these averages are exact for cubics except Hardy's, which is exact only for linear functions. The coefficients in the averages of Table 3 are rational fractions with relatively small denominators, and the user will probably find it convenient to use as weights the integers in the numerators of the coefficients, dividing by the common denominator as the final step. The column headings, therefore, are c_j multiplied by the common denominator.

In both Tables 2 and 3 advantage has been taken of the symmetry of the coefficients c_j to reduce the length of the columns by approximately one-half. The manner of using the tables may be illustrated by taking Spencer's 15-term average as an example. Equation (1.1) shows the calculation of the moving averages. The intermediate values y_x for $x = P - 1$ to $P - 7$ are

calculated successively by the formula

$$y_x = .961572y_{x+1} + .372752y_{x+2} + .015904y_{x+3} - .123488y_{x+4} - .12522$$

$$- .075887y_{x+6} - .025624y_{x+7} .$$

The intermediate values for $x = Q + 1$ to $Q + 7$ are calculated by the identical formula except that the plus signs in the subscripts are changed to minus signs.

The extension procedure drastically reduces the number of values that need to be tabulated for a given weighted average, and makes it possible, for example, to give complete information about 21 such averages in the reasonably compact Tables 2 and 3. However, the user who intends to apply a single weighted average to many data sets may prefer to tabulate the atypical elements of the smoothing matrix G for that weighted average, and so avoid the extra step of calculating the intermediate values. For the benefit of such users, a method of calculating the atypical rows of G will now be described. We observe that the nonzero elements of each row of G except the first m and the last m rows are merely the coefficients c_j of the MWA centered about the diagonal element. The elements in the first m rows of G, except for the first m columns, follow from the symmetry of G, and if $G = (g_{ij})$ we have

$$g_{ij} = c_{j-i} .$$

This leaves only the square submatrix of order m in the upper left corner to be calculated. Let c denote the constant q_{m-s}/p_{m-s}, where p_{m-s} is the term free of z in the polynomial $p(z)$, and let $A_1 = (a_{ij})$ denote the square matrix of order m given by

$$a_{ij} = \begin{cases} 0 & \text{for } i > j \\ 1 & \text{for } i = j \\ -a_{j-i} & \text{for } i < j . \end{cases}$$

Then the required submatrix in the upper left corner of G is given by

$$I - cA_1^T A_1 ,$$

where the superscript T denotes the transpose. The similar submatrix in the lower right corner of G contains the same elements, but with the order of both rows and columns reversed. Justification for this procedure lies in the fact that $F = I - G$ is a symmetric Trench matrix (see the following Section 5).

5. THE GRADUATION MATRIX

In order to describe the unique graduation matrix G of (2.1) that arises when the preferred method of overall graduation is used, it is necessary to define certain special classes of matrices. A square matrix $M = (m_{ij})_{i,j=0}^{N}$ will be called a band matrix if there are nonnegative integers h and k such that $m_{ij} = 0$ whenever $j - i > h$ and also whenever $i - j > k$. Note that we have started the numbering of rows and columns with 0 rather than 1. M will be called strictly banded if, in addition, $h + k \leq N$. In all the banded and strictly banded matrices to be discussed in this paper, h and k will be equal.

M is called a Toeplitz matrix if all the elements on each stripe are equal, where a stripe (Thrall and Tornheim 1957) is either the principal diagonal of the matrix or any diagonal line of elements parallel to the main diagonal. In other words, M is a Toeplitz matrix if there exists a sequence $t_{-N}, t_{-N+1}, \ldots, t_N$ such that for all i and j

$$m_{ij} = t_{j-i} .$$

A strictly banded matrix will be called a Trench matrix if it has a special structure that will now be described. Let $M_i(x)$ denote the generating function of the elements of the ith row: thus,

$$M_i(x) = \sum_{j=0}^{N} m_{ij} x^j .$$

Then M is a Trench matrix if

$$M_i(x) = \begin{cases} A(x)\sum_{j=0}^{i} \beta_j x^{-j} & (0 \leq i < k) \\ A(x)B(1/x) & (k \leq i \leq N-h) \\ B(1/x)\sum_{j=N-i}^{N} \alpha_j x^j & (N-h < i \leq N), \end{cases} \qquad (5.1)$$

where $A(x) = \sum_{j=0}^{h} \alpha_j x^j$ and $B(x) = \sum_{j=0}^{k} \beta_j x^j$ are given polynomials of degree h and k, respectively, with $\alpha_o\beta_o \neq 0$. In all the applications of Trench matrices to be made in this paper, $h = k$ and the coefficients α_j and β_j are real. Careful examination of (5.1) reveals that a Trench matrix is quasi-Toeplitz. By this is meant that the Toeplitz property

$$m_{i+1,j+1} = m_{ij}$$

holds so long as neither of these elements is contained in the k by h submatrix in the upper left corner of M or in the h by k submatrix in the lower right corner. The converse, however, is not true. A strictly banded, quasi-Toeplitz matrix is not necessarily a Trench matrix. In a Trench matrix the elements of the special corner submatrices must be related in a particular way to those of the main part of the matrix as indicated by (5.1).

Greville and Trench (Greville and Trench 1979; Greville 1979, 1980) have studied the properties of Trench matrices. In the joint paper they have shown that a strictly banded matrix has a Toeplitz inverse if and only if it is a nonsingular Trench matrix, and further that a Trench matrix is nonsingular if and only if $A(x)$ and $B(1/x)$ have no common zero.

A rectangular matrix K of $N - s$ rows and N columns will be called a differencing matrix if it transforms a column vector into the column of sth finite differences of the elements of the vector. Evidently the nonzero elements of each row of K are the successive binomial coefficients of order s with alternating signs, and the nonzero elements of K form a diagonal band from the upper left to the lower right.

The following theorem describes the smoothing matrix G of the natural extension.

Theorem 5.1. Let a given MWA of $2m + 1$ terms be exact for the degree $2s - 1$ (with $s \leq m$) and such that $q(z)$ has no zeros on the unit circle and $q_o > 0$. Then, for every $N \geq 2m + 1$, there is a unique square matrix G of order N having the following five properties:

(1) If y is any vector of observed values and $u = Gy$ is regarded as the corresponding vector of graduated values, the elements of u, except the first m and the last m, are merely the graduated values that would be obtained by the use of the given MWA.

(2) G is strictly banded with $h = k = m$.

(3) G is of the form

$$G = I - K^T DK \tag{5.2}$$

for some square matrix D of order $N - s$.

(4) D is nonsingular and has a Toeplitz inverse.

(5) If $D^{-1} = (t_{ij})$, with $t_{ij} = t_{j-i}$, then the series

$$\sum_{\nu=-\infty}^{\infty} t_\nu z^\nu, \tag{5.3}$$

converges in some part of the complex plane.

The unique matrix G so determined has the following properties:

(6) D is unique and G and D are symmetric.

(7) The nonzero elements of the first m and the last m rows of G depend only on the given MWA and do not depend on N. A similar statement applies to the nonzero elements of the first $m - s$ and the last $m - s$ rows of D.

(8) D is a Trench matrix with $B(x) = A(x)$. Moreover,

$$q(x) = A(x)\, A(1/x) \tag{5.4}$$

and the $m - s$ zeros of $A(x)$ are those zeros of $q(x)$ that are outside the unit circle.

(9) $F = I - G = K^T DK$ is a (singular) Trench matrix characterized by the polynomials

$$\hat{A}(x) = \hat{B}(x) = (x-1)^s A(x) \; .$$

(10) The series (5.3) converges to $[q(z)]^{-1}$ in an annulus containing the unit circle.

We shall defer the proof of this theorem until after some explanation and motivation have been given. We begin with a numerical example, which may help to clarify the relationships involved. Then we shall seek to justify the imposition of the five properties that uniquely determine the graduation matrix G. In Section 7 we shall prove the theorem, and in Section 8 we shall show how the extension algorithm can be rationalized and prove that it is mathematically equivalent to the matrix formulation.

For virtually all the MWA's likely to be used in practice the elements of the square submatrices of order m in the upper left and lower right corners of G are irrational. However, for convenience of illustration, we have contrived an example with rational elements. This is the MWA,

$$u_x = \frac{1}{9}\,(2y_{x-2} + y_{x-1} + 3y_x + y_{x+1} + 2y_{x+2}),$$

which is exact only for linear functions and is unlikely to be used in a practical situation. However, it does satisfy Schoenberg's criterion (see Section 11) for a satisfactory smoothing formula. Here $m = 2$, $s = 1$, $q(x) = \frac{1}{9}\,(2x^{-1} + 5 + 2x)$, and $A(x) = \frac{1}{3}\,(2 + x)$. For $N = 7$,

$$G = \frac{1}{9}\begin{bmatrix} 5 & 2 & 2 & 0 & 0 & 0 & 0 \\ 2 & 4 & 1 & 2 & 0 & 0 & 0 \\ 2 & 1 & 3 & 1 & 2 & 0 & 0 \\ 0 & 2 & 1 & 3 & 1 & 2 & 0 \\ 0 & 0 & 2 & 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 2 & 1 & 4 & 2 \\ 0 & 0 & 0 & 0 & 2 & 2 & 5 \end{bmatrix},$$

$$D = \frac{1}{9}\begin{bmatrix} 4 & 2 & 0 & 0 & 0 & 0 \\ 2 & 5 & 2 & 0 & 0 & 0 \\ 0 & 2 & 5 & 2 & 0 & 0 \\ 0 & 0 & 2 & 5 & 2 & 0 \\ 0 & 0 & 0 & 2 & 5 & 2 \\ 0 & 0 & 0 & 0 & 2 & 4 \end{bmatrix},$$

$$T = 3\begin{bmatrix} 1 & -1/2 & 1/4 & -1/8 & 1/16 & -1/32 \\ -1/2 & 1 & -1/2 & 1/4 & -1/8 & 1/16 \\ 1/4 & -1/2 & 1 & -1/2 & 1/4 & -1/8 \\ -1/8 & 1/4 & -1/2 & 1 & -1/2 & 1/4 \\ 1/16 & -1/8 & 1/4 & -1/2 & 1 & -1/2 \\ -1/32 & 1/16 & -1/8 & 1/4 & -1/2 & 1 \end{bmatrix}.$$

Finally,

$$t_\nu = 3(-1/2)^{|\nu|},$$

and it is easily verified that the series (5.3) does in fact converge to $[q(z)]^{-1} = 9(2z^{-1} + 5 + 2z)^{-1}$ for $1/2 < z < 2$. So long as $N \geq 5$, the corresponding results for any value of N could easily be written down.

Let us now look at the five properties that uniquely determine G. The first of these is no more than a restatement of the problem to be solved. The second is a reasonable requirement and amounts to saying that the proposed method of graduation is a "local" procedure: the graduated value of a given observation is not to depend on other observations removed from it by a distance greater than m.

Conditions (3), (4), and (5) are not so obviously appropriate, but are strongly suggested by analogies to the Whittaker graduation method. In order to make these analogies clear, that method is briefly described in the following section.

6. THE WHITTAKER ANALOGIES

The objective of the Whittaker process (Whittaker 1923; Henderson 1924) is to choose graduated values u_j $(j = P, P+1,\ldots,Q)$ in such a way as to minimize the quantity

$$\sum_{j=P}^{Q} W_j(u_j - y_j)^2 + g \sum_{j=P}^{Q-s} (\Delta^s u_j)^2 , \qquad (6.1)$$

where the positive weights W_j , the positive constant g , and the positive integer s are chosen a priori by the user. The solution is most conveniently expressed in matrix notation as follows (Greville 1957, 1974a). Let W denote the diagonal matrix of order N whose successive diagonal elements are the weights W_j , let u and y be defined as in Section 2, and let K be the differencing matrix of Section 5. Then, the expression (6.1) can be written in the form

$$(u - y)^T W(u - y) + g(Ku)^T Ku . \qquad (6.2)$$

It is easily seen (Greville 1974a) that (6.2) is smallest when u satisfies

$$(W + gK^T K)u = Wy . \qquad (6.3)$$

It is not difficult to show (Greville 1957, 1974a) that the matrix in the left member of (6.3) is nonsingular (in fact, positive definite) and therefore

$$u = (W + gK^T K)^{-1} Wy.$$

The Whittaker method has several interesting properties. Commonly the weight W_j is taken as the reciprocal of an estimate of the variance of the jth observation. When this is done, the graduated values are constrained toward the observations where these are reliable, and toward the form of a polynomial of degree $s - 1$ where the observations are less reliable. The method

has the interesting "moment" property

$$\sum_{j=P}^{Q} W_j \, j^{\nu} u_j = \sum_{j=P}^{Q} W_j \, j^{\nu} y_j \qquad (\nu = 0,1,\ldots,s-1).$$

Even in the case of equal weights (equivalent to taking $W = I$) it has been found to be a serviceable method. The ability to choose the constant g at will enables the user to decide how gentle or how drastic he wants the smoothing to be. The remaining discussion will be limited to that case, so that

$$u = (I + gK^T K)^{-1} y.$$

Thus, in terms of (2.1) we have for "unweighted" Whittaker graduation

$$G = (I + gK^T K)^{-1} .$$

It is then easily verified (Noble 1969, page 147) that

$$G = I - K^T(g^{-1} I + KK^T)^{-1} K.$$

This is of the form (5.2) with

$$D = (g^{-1} I + KK^T)^{-1} ,$$

so that property (3) of Theorem 5.1 is also a property of the unweighted Whittaker method. It follows also that

$$D^{-1} = g^{-1} I + KK^T , \tag{6.4}$$

and, if we recall the definition of K in Section 5, it is not difficult to see that KK^T is a Toeplitz matrix. In fact, if $KK^T = (\ell_{ij})$, we have $\ell_{ij} = \ell_{j-i}$, where

$$\ell_{\nu} = (-1)^{\nu} \binom{2s}{s+\nu} ,$$

with the understanding that $\binom{2s}{j}$ vanishes for $j < 0$ and for $j > 2s$. Therefore

$$D^{-1} = (t_{ij}) = t_{j-i} ,$$

where

$$t_{\nu} = g^{-1} \delta_{o\nu} + (-1)^{\nu} \binom{2s}{s+\nu}, \tag{6.5}$$

$\delta_{o\nu}$ being a Kronecker symbol. Thus the unweighted Whittaker method has property (4). Finally, in the Whittaker method D^{-1} given by (6.4) is a band

matrix, and therefore the series (5.3) with t_ν given by (6.5) is finit nd converges everywhere except at the origin. We have, in fact,

$$\sum_{\nu=-\infty}^{\infty} t_\nu z^\nu = g^{-1} + (-1)^s (z^{1/2} - z^{-1/2})^{2s} .$$

As the Whittaker method is a highly regarded graduation procedure, these analogies constitute a strong argument in favor of the natural extension of MWA graduation. Further arguments are provided by the stability theorem of Section 11 and the optimal property of R_o ("reduction of error") for the top and bottom rows of G described in Section 10.

7. PROOF AND DISCUSSION OF THE MAIN THEOREM

Proof of Theorem 5.1. Under the hypotheses of Theorem 5.1, we shall first construct a matrix G having the properties (1) to (10), and shall then show that it is uniquely determined by properties (1) to (5).

As pointed out in Section 3, if ρ is a zero of $q(z)$, it follows from the symmetry of the coefficients that ρ^{-1} is a zero. As there are no zeros on the unit circle, there are $m - s$ zeros inside the unit circle and the same number outside. As any complex zeros must occur in conjugate pairs, there is a polynomial $A(z)$ with real coefficients whose zeros are the $m - s$ zeros of $q(z)$ outside the unit circle. Moreover, the coefficients of $A(z)$ can be normalized so that

$$q(z) = \pm A(z)\, A(1/z) . \tag{7.1}$$

If $A(z) = \sum_{j=0}^{m-s} \alpha_j z^j$, we have

$$q_o = \pm \sum_{j=0}^{m-s} \alpha_j^2 . \tag{7.2}$$

Since, by hypothesis, $q_o > 0$, the positive sign holds in (7.2) and (7.1).

Now, let D be given by property (8): that is, D is the Trench matrix in which the polynomial $A(z)$ of (7.1) plays the role of both $A(x)$ and $B(x)$ in (5.1). It follows from (5.1) that D is symmetric. As the zeros of $A(z)$ are all outside the unit circle, those of $A(1/z)$ are all inside. Thus,

they have no common zero, and it follows (Greville and Trench 1979) that D is nonsingular and D^{-1} is Toeplitz. This verifies property (4), and it follows (see Greville 1979) that the series (5.3) is a "reciprocal" of $q(z)$ at least in the sense that if it is formally multiplied by $q(z)$, the product is unity. It is shown in the paper cited that the series (5.3) converges in some part of the complex plane if (and, under the constraints imposed by (5.4), only if) $A(x)$ is chosen in the manner prescribed by property (8). It is shown further that, when this is done, (5.3) converges to $[q(z)]^{-1}$ in an annulus bounded by circles of radii ρ and ρ^{-1}, where ρ is the minimum absolute value of the zeros of $A(x)$. This verifies properties (5) and (10). It will be noted, incidentally, that the elements of D^{-1} depend only on $A(x)$ and not on N. Increasing N merely extends the sequence $\{t_\nu\}$ without changing the elements already determined.

Now, let G be given by (5.2), so that property (3) holds. The symmetry of G follows from that of D, and property (9) also follows (see Lemma 1 of Greville 1980). Property (7) is a consequence of (5.1) and property (1) follows from (3.2), (5.1), and property (9). Property (2) follows from the fact that F is a Trench matrix.

To prove uniqueness we assume that properties (1) to (5) hold. It then follows from properties (2) and (3) that D is strictly banded with $h = k = m - s$. Because of the structure of K, it is not difficult to see that if D had a nonzero element outside the specified band, G would necessarily have one or more elements outside the band prescribed by property (2). Since, by property (4), D has a Toeplitz inverse, it is a Trench matrix (Greville and Trench 1979).

Let $A(x)$ and $B(x)$ be the polynomials of degree $m - s$ associated with $D = (d_{ij})$, and let

$$d(x) = \sum_{\nu=-m+s}^{m+s} d_\nu x^\nu = A(x)\, B(1/x).$$

Then, by (5.1),

$$d_{ij} = d_{j-i} ,$$

except in the square submatrices of order $m - s$ in the upper left and lower right corners of D. But it follows from (5.2) (see Lemma 1 of Greville 1980) that $F = (f_{ij}) = I - G = K^T DK$ is a (singular) Trench matrix characterized by the polynomials $\hat{A}(x) = (x - 1)^s A(x)$ and $\hat{B}(x) = (x - 1)^s B(x)$. If

$$\hat{d}(x) = \hat{A}(x)\, \hat{B}(1/x) = (-1)^s (x^{1/2} - x^{-1/2})^{2s} d(x) = \sum_{\nu=-m}^{m} \hat{d}_\nu x^\nu ,$$

it follows that

$$f_{ij} = \hat{d}_{j-i}$$

except in the square submatrices of order m in the upper left and lower right corners of F. But it follows from (3.2) and property (1) that, with the possible exception of the elements of the first m and the last m rows of F, f_{ij} is the coefficient of x^{j-i} in $(-1)^s(x^{1/2} - x^{-1/2})^{2s} q(x)$. Therefore $d(x) = q(x)$, and so

$$d_{ij} = q_{j-i} \tag{7.3}$$

except in the first $m - s$ and the last $m - s$ rows of D, and

$$q(x) = A(x)\, B(1/x) . \tag{7.4}$$

Now, (7.3) and the fact that D is a Trench matrix (and has the quasi-Toeplitz property) uniquely determine all its elements except those of the two special corner submatrices of order $m - s$. It has been shown elsewhere (Greville 1979) that, in general, these corner elements can be chosen in a (finite) number of ways, corresponding to the possible ways of factoring the polynomial $x^{m-s} q(x)$ of degree $2m - 2s$ as the product of two polynomials, $A(x)$ and $x^{m-s} B(1/x)$ of degree $m - s$ with real coefficients. (Two factorizations are considered different only if the set of zeros of $A(x)$ is different.) Property (8) prescribes a unique factorization, and therefore G is uniquely determined by properties (1) to (4) and (8). But it has been shown (Greville 1979) that if $q(x)$ is given and (7.3) and (7.4) hold,

properties (8) and (5) are equivalent. Thus, G is uniquely determined by properties (1) to (5). ∎

At the beginning of Theorem 5.1 two general hypotheses concerning the given MWA were introduced: that q_o is positive, and that $q(z)$ has no zeros on the unit circle. These will now be discussed. It follows from (7.2) that $q_o \neq 0$; it is either positive or negative. Typically, the coefficients q_j are all positive and q_o is the largest. No reasonable MWA could have q_o negative. In fact, it will be shown in Section 11 that an MWA with negative q_o cannot satisfy Schoenberg's criterion for a satisfactory smoothing formula.

MWA's with zeros of $q(z)$ on the unit circle are also uncommon (I know of no example in the published literature), but it is easy to construct such formulas, and the presence of such zeros does not of itself suggest any inherent defect in the MWA. However, when there are such zeros, property (5) of Theorem 5.1 cannot hold (see Greville 1979), and no single preferred method of extending the graduation to the ends of the data is clearly indicated. The case of zeros of $q(z)$ on the unit circle is of no practical importance, and its inclusion in Theorem 5.1 would have made the theorem messy and complicated. The curious reader might wish to consult Greville's (1979, 1980) Manitoba Conference papers for discussion of the mathematical questions involved.

8. DEVELOPMENT OF THE NATURAL METHOD THROUGH THE EXTENSION ALGORITHM

This section has two objectives: first, to show that the extension algorithm of Section 3 is mathematically equivalent to the matrix approach of Sections 5 and 7, and second, to provide a heuristic argument for the natural extension of the graduation that is a partial alternative to the five determining properties listed in Theorem 5.1.

Because of the symmetry of the coefficients q_j of $q(z)$, and the fact that it has no zeros on the unit circle, there is, if $q_o > 0$, a polynomial

$A(z)$ of degree $m - s$ with real coefficients such that (5.4) holds. If $m > s$, there is more than one such polynomial. Let the polynomial

$$\hat{A}(z) = \sum_{j=0}^{m} \hat{\alpha}_j z^j \tag{8.1}$$

of degree m be defined by

$$\hat{A}(z) = (z - 1)^s A(z) . \tag{8.2}$$

Then it follows from (3.2) that

$$u_x = y_x - \hat{A}(E) \hat{A}(E^{-1}) y_x . \tag{8.3}$$

Thus it would be possible to perform the graduation by means of the following three steps:

1. Operate on the sequence $\{y_x\}$ with $\hat{A}(E^{-1})$ to obtain the sequence of values of $\hat{A}(E^{-1}) y_x$.

2. Operate on the latter with $\hat{A}(E)$ to obtain the sequence of values of $\hat{A}(E) \hat{A}(E^{-1}) y_x$, which may be thought of as corrections to the observed values y_x.

3. Subtract each of the latter values from the corresponding value of y_x to obtain the graduated values u_x.

Now, suppose the sequence $\{y_x\}$ is extended backward to $x = P - m$ and forward to $x = Q + m$ by imposing the conditions

$$\hat{A}(E) y_x = 0 \qquad (x = P - 1, P - 2, \ldots, P - m) \tag{8.4}$$

and

$$\hat{A}(E^{-1}) y_x = 0 \qquad (x = Q + 1, Q + 2, \ldots, Q + m). \tag{8.5}$$

These extensions make it possible to calculate by the main formula graduated values u_x over the entire range $x = P, P + 1, \ldots, Q$. For $x = P, P + 1, \ldots, P + m - 1$, (8.1) and (8.3) give

$$y_x - u_x = \sum_{j=0}^{m} \hat{\alpha}_j \hat{A}(E) y_{x-j} .$$

In view of (8.4) this reduces to

$$y_x - u_x = \sum_{j=0}^{x-P} \hat{\alpha}_j \hat{A}(E) y_{x-j} . \tag{8.6}$$

A similar argument applies for $x = Q, Q-1, \ldots, Q-m+1$, with $\hat{A}(E)$ and $\hat{A}(E^{-1})$ interchanged, and we have there

$$y_x - u_x = \sum_{j=0}^{Q-x} \hat{\alpha}_j \hat{A}(E^{-1}) y_{x+j} . \tag{8.7}$$

If we write $F = (f_{ij})$, it follows from (8.6) and (8.7) that

$$y - u = Fy ,$$

where F is the (singular) Trench matrix characterized by two polynomials of degree m both equal to $\hat{A}(x)$. But in view of property (8) of Theorem 5.1, this is exactly the matrix $F = I - G$ of property (9) of that theorem. Thus the two approaches give the same result if the polynomial $A(x)$ is the same in both cases. Therefore let us consider the choice of $A(x)$ in the present context.

For a moment let us think of the sequence $\{y_x\}$ as extended indefinitely to the left of $x = P$ rather than only as far as $x = P - m$. Then, the general solution of (8.4)

$$y_x = P_1(x) + \sum_{j=1}^{m-s} b_j r_j^x , \tag{8.8}$$

where $r_1, r_2, \ldots, r_{m-s}$ are the zeros of $A(x)$, $P_1(x)$ is an arbitrary polynomial of degree less than s, and $b_1, b_2, \ldots, b_{m-s}$ are arbitrary constants. If $A(x)$ has multiple zeros, (8.8) is replaced by a slightly different expression, but the end result is the same.

Similarly, the general solution of (8.5) is

$$y_x = P_2(x) + \sum_{j=1}^{m-s} e_j r_j^{-x} . \tag{8.9}$$

Let us now impose the conditions that $\Delta^s y_x$ with y_x given by (8.8) shall approach zero as x tends to $-\infty$, and that $\Delta^s y_x$ with y_x given by (8.9) shall approach zero as x tends to ∞. Since $A(x)$ must satisfy (5.4), it is clear that these conditions are satisfied if and only if $A(x)$ is chosen so that its zeros are the $m - s$ zeros of $q(x)$ that are outside the unit circle. Clearly, this is the choice of $A(x)$ prescribed by Theorem 5.1. Moreover,

$A(x)$ chosen in this manner is closely related to the polynomial $p(x)$ of Section 3. In fact,

$$A(x) = \alpha_{m-s} x^{m-s} p(1/x),$$

and consequently,

$$\hat{A}(x) = \alpha_{m-s} x^{m-s} a(1/x) ,$$

where $a(x)$ is defined by (3.4). Thus, with this choice of $A(x)$, (8.4) and (8.5) are equivalent to the extension algorithm of Section 3.

We note in passing that the computational short cut involving extended values has an analogue in the case of Whittaker smoothing. Especially in actuarial literature, the Whittaker smoothing process is sometimes called the difference-equation method because the difference equation

$$u_x + (-1)^s \delta^{2s} u_x = y_x \tag{8.10}$$

holds for $x = P + s, P + s + 1, \ldots, Q - s$. Aitken (1926) pointed out that (8.10) is satisfied for $x = P, P + 1, \ldots, Q$ if we annex at each end of the data set s extrapolated values of both y_x and u_x satisfying the conditions

$$u_x = y_x \qquad (x = P - j,\ x = Q + j;\ j = 1, 2, \ldots, s),$$
$$\Delta^s u_x = 0 \qquad (x = P - j,\ x = Q - j;\ j = 1, 2, \ldots, s).$$

However, this observation is not helpful from a computational point of view. The attempt to utilize it merely increases from N to $N + 2s$ the order of the linear system to be solved.

A final comment regarding motivation of the extension algorithm may be in order. Elsewhere (Greville 1957) I referred to the notion of the "smooth space." This is the space of vectors y such that $Gy = y$, the space of vectors that are unchanged by the graduation process. For the Whittaker process it is the space of "polynomial vectors" of degree less than s .

In a strict mathematical sense, the smooth space is the same for the graduation procedure considered here, but in reality the situation is more complicated. Equation (8.3) shows that if (8.4) should hold "across the board," all the graduated values, to the extent that they can be calculated by the main formula without extension, would be equal to the observed values. Of course, the result would be similar if (8.5) should hold "across the board." Now, the two conditions (8.4) and (8.5) are not equivalent. The corresponding conditions in the Whittaker case are the vanishing of the sth finite differences of the observed values, which are equivalent because of the symmetry of the (binomial) coefficients in the expressions for these finite differences. The coefficients of $\hat{A}(x)$ have no such symmetry.

These observations suggest that the true analogue of the Whittaker process is arrived at by using the different criteria (8.4) and (8.5) at the two ends of the data. As previously noted, there are in general different ways of choosing $A(x)$ so that

$$\hat{A}(E)\ \hat{A}(E^{-1}) = (-1)^s\ \delta^{2s}\ q(E)\ ,$$

and we have made the unique choice that makes the extension a "stable" operation at both ends.

9. SPECIAL CLASSES OF MOVING AVERAGES

Of particular interest are those moving averages known to actuaries as minimum-R_3 formulas and to economic statisticians as "Henderson's ideal" formulas. For a given number of terms $2m + 1$, this is the average (1.2), exact for the third degree, for which the quantity

$$\sum_{j=-m-3}^{m} (\Delta^3\ c_j)^2 \tag{9.1}$$

is smallest (with the understanding that $c_j = 0$ for $|j| > m$). The "smoothing coefficient" R_3 is defined as the quantity obtained by dividing (9.1) by 20 and taking the square root. The divisor 20 is chosen because this is the

value of (9.1) for the trivial case of (1.2) in which $c_0 = 1$ and $c_j = 0$ for $j \neq 0$.

The rationale for minimizing (9.1) may be explained as follows (Greville 1974a). If, for some x, u_x, u_{x+1}, u_{x+2}, and u_{x+3} are all given by (1.2) (which is the case for $x = P + m$ to $Q - m - 3$, inclusive), then

$$\Delta^3 u_x = - \sum_{j=-m-3}^{m} (\Delta^3 c_j) y_{x+j+3} . \tag{9.2}$$

It has been customary to regard the smallness (in absolute value) of the third differences of the graduated values as an indication of smoothness. Therefore (9.2) suggests that smoothness is encouraged by making the quantities $\Delta^3 c_j$ numerically small, and minimizing (9.1) is a way of doing this. The formula corresponding to (9.2) for a general order of differences is

$$\Delta^s u_x = (-1)^s \sum_{j=-m-s}^{m} (\Delta^s c_j) y_{x+j+s}, \tag{9.3}$$

and the general formula for R_s is

$$R_s^2 = \sum_{j=-m-s}^{m} (\Delta^s c_j)^2 / \binom{2s}{s}. \tag{9.4}$$

There is some question whether Henderson's contribution warrants attaching his name to the "ideal" weighted averages. De Forest (1873) treated extensively the formulas that minimize R_4. The concept of choosing the coefficients c_j in order to minimize R_3 seems to have been first mentioned by G. F. Hardy (1909). These averages were fully discussed by Sheppard (1913) slightly earlier than by Henderson (1916). However, Henderson does seem to have been the first to give an explicit formula for the coefficient c_j in the weighted average minimizing R_3 (Henderson 1916, p. 43; Macaulay 1931, p. 54; Henderson 1938, p. 60; Miller 1946, p. 71; Greville 1974a, p. 18). If we write $k = m + 2$, so that the weighted average has $2k - 3$ terms, the formula is

$$c_j = \frac{315[(k-1)^2 - j^2](k^2 - j^2)[(k+1)^2 - j^2](3k^2 - 16 - 11j^2)}{8k(k^2 - 1)(4k^2 - 1)(4k^2 - 9)(4k^2 - 25)} . \tag{9.5}$$

Weighted averages that minimize R_s have been discussed from other points of view by Wolfenden (1925), Schoenberg (1946), and Greville (1966, 1974b).

Also deserving of special mention are the averages (exact for cubics) that minimize R_o, sometimes called "formulas of maximum weight" or "Sheppard's ideal" formulas. These are sometimes applied to physical measurements when the errors of observation can be regarded as random "white noise" (see discussion of "reduction of error" in Section 10). The weights are given by

$$c_j = \frac{3(3m^2 + 3m - 1) - 15j^2}{(2m - 1)(2m + 1)(2m + 3)} .$$

Weighting coefficients c_j and extension coefficients a_j for minimum-R_3 (Henderson's ideal) averages of 5, 7, ..., 23 terms are given in Table 2.

10. COMPARISON WITH OTHER METHODS. PRACTICAL CONSIDERATIONS

If a symmetrical MWA exact for the degree $2s - 1$ is being used to smooth the main part of the data, it can easily be deduced, either from the extension algorithm described in Sections 3 and 8 or from the matrix formulation of Theorem 5.1 that the unsymmetrical weightings proposed for smoothing the first m and the last m observations are exact only for the degree $s - 1$. For example, all the averages represented in Tables 2 and 3 with the exception of Hardy's are exact for cubics, and therefore their extensions to the ends are exact only for linear functions. Hardy's weighted average is exact for linear functions and its extension only for constants.

The Whittaker process has a *similar* property. At a sufficient distance from the ends of the data, polynomials of degree $2s - 1$ are "almost" reproduced by that process. In support of this rather loose statement the following heuristic argument is advanced. For the Whittaker process

$$G = (I + gK^T K)^{-1} = I - gGK^T K.$$

Thus, if y is the vector of observed values, the vector of corrections to these values is

$$-gGK^T Ky.$$

Now, the nonzero elements of $K^T K$, with the exception of the first s and the last s rows, are binomial coefficients of order $2s$ with alternating signs. Therefore the components of $K^T Ky$, except for the first s and the last s, are (2s)th differences of those of y (or their negatives if s is odd). Thus, if y is a vector of equally spaced ordinates of a polynomial of degree $2s - 1$, $K^T Ky$ is a vector of zeros except for the first s and the last s components. The components of $GK^T Ky$ are graduated values of those of $K^T Ky$, and therefore should be very small at some distance from the extremities of the data. Finally, multiplication by g, even though g is typically large, should give small corrections at a sufficient distance from the ends of the data.

Some users may consider the reduction in degree of exactness near the ends a disadvantage of the natural method of extension. Before I became aware of the natural method, I had proposed (Greville 1974a) a different method of extension (already mentioned in Section 2) that does not have this particular disadvantage (though it has other shortcomings). This involves extrapolation by a polynomial of degree $2s - 1$ fitted by least squares to the first $m + 1$ observations. A similar polynomial is fitted to the last $m + 1$ observations for extrapolation at the other end of the data. There may be a gain in simplicity by using a single method of extrapolation for all symmetrical weighted averages, so that the extrapolated values depend only on the number of terms in the main formula. However, there is a loss in that the extension method is no longer tailored to the particular symmetrical average used.

Like the natural method of extension, the method using extrapolation by least squares can be collapsed into a single matrix G. When this is done, the band character of the smoothing matrix is maintained, but the symmetry is lost. Though the matrix approach is less convenient for computational

purposes, the differences between the two methods are best elucidated by comparing the first m rows of the respective matrices G . This is done in Tables 4 and 5 for the case of the 9-term "ideal" formula. Here m = 4, but for convenience the fifth row is also shown. Its elements would be repeated in the subsequent rows, moving successively to the right, until we come to the last four rows. While an average of as few as 9 terms would seldom be used in practice, this is a convenient illustration.

As previously indicated, the first m and the last m rows of G may be regarded as exhibiting unsymmetrical weighted averages which are to be used near the ends of the data to supplement the symmetrical average used elsewhere. The coefficients that appear in the last m rows are the same as those in the first m rows, but the order is reversed, both horizontally and vertically. It should be noted that the coefficients in the supplemental averages depend only on those of the underlying symmetrical average. They do not depend on N , the number of observations in the data set (which is the order of G).

The coefficients in the supplemental weighted averages based on least-squares extrapolation, exhibited in Table 5, show two undesirable features. These are negative coefficients of substantial numerical magnitude, and successive waves of positive and negative coefficients as one proceeds from left to right along the rows. The number of such waves would increase as the number of terms in the underlying formula increases.

In striking contrast is the character of the coefficients of the natural extension. Like the coefficients in the underlying symmetrical formula, each row exhibits a peak in the vicinity of the main diagonal of the matrix, tapering off to a single group of negative coefficients of reduced size near the edge of the band.

In the least-squares method only a very small correction is made to the initial observed value. The corresponding correction in the natural method is more substantial.

The "second-difference correction" is the coefficient of the second-difference term when the formula is expressed in terms of increasing orders of differences in the form

$$u_x = y_x + c\Delta^2 y_{x-h} + \cdots .$$

The coefficient c does not depend on the subscript $x - h$, in which there is some freedom of choice. For the formulas based on least-squares extrapolation, which are exact for cubics, the fourth-difference correction is similarly defined.

Some writers (Miller 1946; Wolfenden 1942; Greville 1974a) have regarded the observed values y_x as the sum of "true" values U_x and superimposed random errors e_x. If it is assumed that the errors e_x for different x are uncorrelated, and have zero mean and constant variance σ^2 for all x, then the variance of the error in the smoothed value u_x is $R_o^2 \sigma^2$, where R_o^2 is obtained by taking $s = 0$ in (9.4). Thus, R_o may be interpreted as the ratio of reduction in the standard deviation of error that results from application of the weighted average.

While the assumptions underlying the preceding analysis may be questioned, nevertheless a good case can be made that, for any weighted average, R_o should be less than unity. Since R_o^2 is the sum of the squares of the coefficients in the average, R_o can never be less than the maximum of the absolute values of the coefficients. Thus, an average cannot be considered satisfactory if the absolute value of any coefficient is equal to or greater than unity.

When the graduation is extended to the extremities of the data, these remarks apply not only to the main formula but also to the unsymmetrical formulas to be used near the ends. Tables 4 and 5 illustrate the fact that there is a strong tendency for R_o to become large as we approach the extremities of the data.

In this connection, the natural extension has an important optimal property. Let us suppose that the main formula is given and satisfies the general hypotheses of Theorem 5.1. That is, it is symmetrical, $q_o > 0$, and $q(z)$ has no zeros on the unit circle. Further we suppose that $c_o > -1$. The latter assumption is not a strong one; a negative value of c_o is most unusual in any case, and we have previously stated that an average is not satisfactory if the absolute value of any coefficient is equal to or greater than unity. In addition we assume that G is symmetric and has properties (1) to (4) of Theorem 5.1. As we shall see in Section 11, there are cogent reasons for thinking that G should be symmetric.

Under these conditions we have seen that in general there are a number of possible choices of the polynomial $A(x)$. It will be shown that R_o for the top row of G is smallest when $A(x)$ is chosen in the unique manner prescribed by Theorem 5.1.

Since G is symmetric and $q_o > 0$, the coefficients of $A(x)$ and $B(x)$ can be normalized so that $B(x) = A(x)$. Using the notation of Sections 7 and 8, let

$$\hat{A}(x) = (x - 1)^s A(x) = \sum_{j=0}^{m} \hat{a}_j x^j .$$

Then, the middle weight of the MWA is one minus the constant term in the expansion of $\hat{A}(x) \hat{A}(1/x)$, or in other words,

$$c_o = 1 - \sum_{j=0}^{m} \hat{a}_j^2 . \tag{10.1}$$

Now, since $\hat{\beta}_o = \hat{\alpha}_o$, the nonzero elements in the top row of G are, successively, $1 - \hat{\alpha}_o^2$, $-\hat{\alpha}_o \hat{\alpha}_1$, $-\hat{\alpha}_0 \hat{\alpha}_2$, ..., $-\hat{\alpha}_o \hat{\alpha}_m$. Therefore, R_o^2 for this top row is given by

$$R_o^2 = 1 - 2\hat{\alpha}_o^2 + \hat{\alpha}_o^2 S = 1 - \hat{\alpha}_o^2(1 + c_o), \qquad (10.2)$$

where S denotes the summation contained in (10.1). Now, let $r_1, r_2, \ldots, r_{m-s}$ denote the zeros of $A(x)$, so that

$$A(x) = \alpha_{m-s} \prod_{j=1}^{m-s} (x - r_j) ,$$

and let

$$\lambda = (-1)^{m+1} \prod_{j=1}^{m-s} r_j .$$

Then,

$$\hat{\alpha}_o = -\lambda\alpha_{m-s} , \qquad \hat{\alpha}_m = \alpha_{m-s} ,$$

$$c_m = -\hat{\alpha}_o \hat{\alpha}_m = \lambda\alpha_{m-s}^2 , \qquad \hat{\alpha}_o^2 = \lambda^2 \alpha_{m-s}^2 ,$$

so that $\hat{\alpha}_o^2 = \lambda c_m$, and (10.2) becomes

$$R_o^2 = 1 - \lambda c_m(1 + c_o) .$$

In this expression, c_o and c_m are given; λ is the only variable. Moreover, $\lambda c_m = \hat{\alpha}_o^2$ is positive, and $1 + c_o$ is positive, since $c_o > -1$. Therefore R_o^2 is smallest when $|\lambda|$ is largest, which is clearly the case when the zeros of $A(x)$ are the $m - s$ zeros of $q(x)$ that are largest in absolute value, namely, those outside the unit circle.

Thus the smoothing matrix G of the natural extension would still be uniquely determined if, in Theorem 5.1, we replaced property (5) by the requirements that G be symmetric and that R_o for the top row of G be as small as possible subject to the other conditions imposed. It appears that the requirement that G be symmetric can be dropped if stronger conditions are imposed on $q(z)$, but the algebraic complication of the proof is greatly increased.

As indicated in Section 9, it has long been customary to regard a graduation as smooth if the third differences of the graduated values are small in

absolute value. If $G = (g_{ij})$, we have

$$u_{P+i-1} = \sum_{j=1}^{N} g_{ij}\, y_{P+j-1} \,,$$

and therefore

$$\Delta^s\, u_{P+i-1} = \sum_{j=1}^{N} y_{P+j-1}\, \Delta_i^s\, g_{ij} \,, \qquad (10.3)$$

where the subscript of Δ indicates that the differences are taken with respect to i (i.e., down the columns of the matrix). If one avoids the corner submatrices, the nonzero elements g_{ij} in (10.3) are merely coefficients in the underlying symmetrical average, and (10.3) reduces to (9.3). This was the rationale underlying the derivation of the minimum-R_s averages.

Of course, if G is symmetric, it makes no difference whether the differences are taken horizontally or vertically. When the symmetry of G is not assumed, care must be exercised. Many years ago (Greville 1947, 1948) I published what purported to be coefficients in supplemental averages to be used near the ends of the data in conjunction with minimum-R_3 and minimum-R_4 symmetrical averages. The symmetry of G was not assumed, and I made the error of deriving the unsymmetrical coefficients by minimizing their third and fourth differences taken horizontally. The tables in question are therefore based on an incorrect assumption. Further it may be mentioned in passing that in the 1947-8 formulation the diagonal band character was not maintained, since the supplemental averages contained the full $2m + 1$ terms.

Table 6 shows, for the natural and least-squares extensions of the 9-term minimum-R_3 formula, those third differences of the matrix elements, taken vertically, that involve elements of the first five rows. The entries in the fifth row of Table 6 would be repeated in subsequent rows, moving successively to the right. Casual inspection of the table shows that the third differences are numerically smaller for the natural extension. All of these third differences are less than 0.14 in absolute value. Two of those for the least-squares

extension exceed 0.7 in absolute value.

It is instructive to compare the natural extension with the least-squares extension for the numerical example of Section 3. Though neither extension is recommended for use when additional data are available beyond the range of the original data set, nevertheless it may be of interest, purely for purposes of illustration, to choose a numerical example in which such additional data are available, and this has been done.

Table 7 and Figures B and C complement Table 1 and Figure A, showing, for the first seven months of 1967 and the last seven months of 1971, the observed values of precipitation in Madison, Wisconsin, and the graduated values obtained by (i) natural extension of Spencer's 15-term average, (ii) least-squares extension of the same average, and (iii) use of additional data. It will be noted that the least-squares extension is strongly constrained toward each of the two terminal observations (January 1967 and December 1971). This may be explained by the fact that all the values y_x in (1.2) that entered into the calculation of these graduated values are included in either the $m + 1$ observations to which the least-squares cubic was fitted or the m extrapolated values obtained from the same cubic. On the other hand, the natural extension and the least-squares extension are very close together at the interface with the graduated values calculated in the standard manner. Thus, for the months of July 1967 and June 1971, all but one of the values y_x entering into the computation (1.2) are identical for the two methods.

For the months closer to the interface the graduated values obtained by introducing additional data are close to those of the natural extension. This is because the supplemental unsymmetrical averages produced by the natural extension (unlike those of the least-squares extension) give relatively small

weight to the observations more remote from the one being graduated (as does the underlying symmetrical formula). For example, the values for the natural extension and those obtained by the use of additional data are indistinguishable in Figure B for April to July 1967. In the last months of 1971 the deviation is greater because the first two months of 1972 were exceptionally dry. This could not have been predicted from the data for preceding months.

Table 8 gives certain parameters for the various symmetrical weighted averages that have been mentioned previously. The column headed "Error" requires explanation. This is the error committed when the formula in question is used to "smooth" a polynomial of degree 4. This naturally tends to increase with the number of terms in the formula. Both R_o and R_3 tend to decrease with increasing number of terms. Though the "ideal" formulas have been derived to minimize R_3, they tend to produce small values of R_o as well. In only one instance (Vaughan) does a "name" formula have a smaller R_o than the "ideal" formula of the same number of terms. The late Hubert Vaughan was a remarkably keen analyst of MWA smoothing.

It may be mentioned in passing that some writers (e.g., Henderson 1938) call the reciprocal of R_o^2 the "weight" and the reciprocal of R_3 the (smoothing) "power."

11. THE STABILITY THEOREM

Schoenberg (1946) defined the <u>characteristic function</u> of the MWA (1.2) as

$$\phi(t) = \sum_{j=-m}^{m} c_j e^{ijt} . \qquad (11.1)$$

For a symmetrical MWA this is a real function of the real variable t, and can be expressed in the alternative form

$$\phi(t) = \sum_{j=-m}^{m} c_j \cos jt .$$

It is periodic with period 2π and equal to unity for $t = 2\pi n$ for all

integers n .

The effect of MWA's in eliminating or reducing certain waves has been noted (Elphinstone 1951; Hannan 1970). If the input to the smoothing process is a sine wave, which may be represented in the form

$$y_x = C \cos(rx + h) \ , \qquad (11.2)$$

it can be shown by simple algebraic manipulation that

$$u_x = y_x \ \phi(2\pi/\omega) \ ,$$

where $\omega = 2\pi/r$ is the period of y_x . Thus, if $\phi(2\pi/\omega) = 0$, the wave is annihilated by the smoothing process; the amplitude is severely reduced if it is close to zero. Thus MWA smoothing is related to the "filtering" processes considered by Wiener (1949) and others.

Schoenberg (1946) defined a <u>smoothing formula</u> as an MWA whose characteristic function $\phi(t)$ satisfies the condition

$$|\phi(t)| \leq 1 \qquad (11.3)$$

for all t. Thomée (1965) calls (11.3) "von Neumann's condition" without, however, citing any specific publication of von Neumann. Later Schoenberg (1948, 1953) suggested the stronger condition

$$|\phi(t)| < 1 \qquad (0 < t < 2\pi). \qquad (11.4)$$

Lanczos (see Schoenberg 1953) pointed out that (11.4) is obtained by requiring that every simple vibration (11.2) be diminished in amplitude by the transformation (1.2). The results of Section 5 of the present paper suggest an alternative definition of a smoothing formula. Using the subscript N to emphasize the fact that the order of G is the number of observations in the data set, we shall say that an extension of (1.2) by means of a smoothing matrix G is <u>stable</u> if the limit

$$G_N^\infty = \lim_{n \to \infty} G_N^n$$

exists for all N . Schoenberg (1953, footnote 3) suggested a relationship

between (11.3) and the conditions for existence of the infinite power of a matrix (Oldenburger 1940; Dresden 1942), but he did not elaborate the connection. In the theorem of this section we shall attempt to do so.

In Section 7 we promised to justify the hypothesis in Theorem 5.1 that $q_o > 0$ by showing that if the characteristic function of a symmetric MWA satisfies (11.3) and $q(z)$ has no zeros on the unit circle, then $q_o > 0$. Consider the real function

$$\psi(t) = 1 - \phi(t) \tag{11.5}$$

and note that (11.3) is equivalent to

$$0 \leq \psi(t) \leq 2 \tag{11.6}$$

for all t. From (3.1), (3.2), and (11.1) it follows that

$$\psi(t) = (-1)^s (2i \sin \tfrac{1}{2}t)^{2s} q(e^{it}) = (4\sin^2 \tfrac{1}{2}t)^s \, q(e^{it}), \tag{11.7}$$

and therefore (11.6) implies that $q(e^{it})$ *is nonnegative for* $0 < t < 2\pi$. In fact, it is positive, since $q(z)$ has no zeros on the unit circle, and by continuity it is positive for $t = 0$ as well. In other words, $q(1) > 0$. Now let the polynomials $A(x)$ and $B(x)$ be chosen so that $q(x) = A(x)\,B(1/x)$ and the zeros of $B(1/x)$ are the reciprocals of those of $A(x)$. This is always possible because of the symmetry of the coefficients of $q(x)$. Moreover, the coefficients in those polynomials can be normalized, as in the proof of Theorem 5.1, so that (7.1) and (7.2) hold, and therefore

$$q(1) = \pm[A(1)]^2.$$

Since $q(1)$ is positive, the positive sign holds in (7.1) and (7.2), and consequently $q_o > 0$.

Before stating the theorem that elucidates the relationship of condition (11.3) to the smoothing matrix G, we need to describe certain results published elsewhere that will be used in the proof. In a recent paper (Greville 1980) I have studied bounds for eigenvalues of Hermitian Trench matrices

(which become symmetric Trench matrices when the elements are real). If the polynomials that characterize a real symmetric Trench matrix H are $A(x)$ and $B(x)$ of degree h, we have seen that the coefficients can be normalized so that either $B(x) = A(x)$ or $B(x) = -A(x)$. If the minus sign holds, one can consider the symmetric Trench matrix $-H$. It is sufficient, therefore, to consider the case in which $B(x) = A(x)$.

Let $A(x)$ be given and consider the family of symmetric Trench matrices H_N of order N $(N \geq 2h + 1)$ characterized by $A(x)$ and $B(x) = A(x)$. Let

$$G_N = I - \mu H_N ,$$

where μ is a positive constant, and let

$$h(x) = A(x)\, A(1/x).$$

Then it is shown that $h(x)$ is real and nonnegative on the unit circle, and has a maximum thereon, which we denote by M. Then Corollary 1 of the cited paper states that the limit G_N^{∞} exists for all N if and only if

$$\mu \leq 2/M$$

and no zero of $A(x)$ is inside the unit circle unless it is also a zero of $A(1/x)$. A particular application of Lemma 1 of the same paper yields the result that if D is a Trench matrix characterized by the polynomials $A(x)$ and $B(x)$, then $K^T DK$ (with K defined as in Section 5) is a (singular) Trench matrix characterized by the polynomials $\hat{A}(x) = (x - 1)^s A(x)$ and $\hat{B}(x) = (x - 1)^s B(x)$. For convenience in the proof of the theorem that follows, we shall refer to Corollary 1 and Lemma 1 of the paper cited as merely "Corollary 1" and "Lemma 1."

<u>Theorem 11.1.</u> Let a symmetrical MWA (1.2) be given and let the associated smoothing matrix G_N for all $N \geq 2m + 1$ be symmetric and have properties (1) to (4) of Theorem 5.1. Then the family of matrices G_N is stable if and only if (11.3) holds and the polynomial $A(x)$ associated with the matrix D has no zero inside the unit circle.

Proof. From the hypotheses stated in the first sentence of the theorem we can deduce certain properties of the matrices F and D by reasoning similar to that used in the uniqueness part of the proof of Theorem 5.1. First we note that the hypotheses of the present theorem differ slightly from those of Theorem 5.1. We have added the hypothesis that G is symmetric, and have omitted the restrictions on $q(x)$. However, the reader should note carefully that the latter omission is occasioned only by the fact that these restrictions are implied by the symmetry of G in conjunction with other hypotheses. The symmetry of G implies that of F. As the rows of K are linearly independent, it has full row rank and therefore has a left inverse, say L (see Ben-Israel and Greville 1974, Lemma 1.2). Therefore,

$$L^T FL = L^T K^T DKL = D ,$$

and consequently D is symmetric.

As in the uniqueness part of the proof of Theorem 5.1, it follows from property (4) that D is a nonsingular Trench matrix. If it is characterized by the two polynomials $A(x)$ and $B(x)$ of degree $m - s$, then

$$q(x) = A(x)\, B(1/x) ,$$

as in the earlier proof. As D is real and symmetric, the coefficients in these polynomials are real and can be normalized so that

$$B(x) = \pm A(x). \tag{11.8}$$

As we have omitted the hypothesis that $q_o > 0$, some ambiguity remains about the sign of the right member of (11.8) until further hypotheses are introduced, and we have

$$q(x) = \pm A(x)\, A(1/x). \tag{11.9}$$

Now, the symmetry and nonsingularity of D and the requirement that $A(x)$ have real coefficients imply that $q(x)$ has no zeros on the unit circle. As we have seen, symmetry of D implies (11.9), and nonsingularity implies

(Greville and Trench 1979) that $A(x)$ and $A(1/x)$ have no common zero. Now, if $q(x)$ has a zero on the unit circle, say x_o, then x_o^{-1} is also on the unit circle, so that $A(x)$ must have a zero on the unit circle; call it ρ. Then ρ^{-1} is a zero of $A(1/x)$, and $\bar{\rho}$ is a zero of $A(x)$, since $A(x)$ has real coefficients. But $\bar{\rho} = \rho^{-1}$; therefore $A(x)$ and $A(1/x)$ have a common zero. Thus the supposition that $q(x)$ has a zero on the unit circle is false.

Now, suppose that (11.3) holds and $A(x)$ has no zeros inside the unit circle. Then it follows from the discussion following (11.7) that the positive sign holds in (7.1) and (7.2), and therefore in (11.9). By Lemma 1, F is a singular Trench matrix characterized by the polynomials

$$\hat{A}(x) = \hat{B}(x) = (x - 1)^s A(x). \tag{11.10}$$

Let

$$f(x) = \hat{A}(x)\, \hat{B}(1/x) = (x^{1/2} - x^{-1/2})^{2s}\, q(x).$$

Then,

$$\psi(t) = f(e^{it}). \tag{11.11}$$

Let M denote the maximum value of $f(x)$ on the unit circle. Then by (11.6) $M \leq 2$, or

$$1 \leq 2/M. \tag{11.12}$$

Consequently, by Corollary 1, the family of matrices G_N is stable.

Conversely, suppose that the family $\{G_N\}$ is stable, in addition to the hypotheses in the first sentence of the theorem. Since G_N is symmetric, its eigenvalues are real, and stability implies (Oldenburger 1940; Dresden 1942) that all its eigenvalues are in the half-open interval $(-1,1]$. In other words, all the eigenvalues of F_N are in $[0,2)$ for all N. Now, if v is an arbitrary column vector of real elements, it is well known that the minimum value of the Rayleigh quotient $v^T Fv/v^T v$ is the (algebraically)

smallest eigenvalue of F. Suppose the minus sign holds in (11.8) and let v be the unit vector with 1 as its first element and all the other elements 0. By (11.10), the constant term of $\hat{A}(x)$ is $(-1)^s \alpha_o$, and the Rayleigh quotient is $-\alpha_o^2$, which is negative since $\alpha_o \neq 0$ by the definition of a Trench matrix. Thus, F has a negative eigenvalue, in contradiction to the statement that all its eigenvalues are in $[0,2)$. Therefore the supposition that the minus sign holds in (11.8) is false.

Since the positive sign holds in (11.8), F belongs to the class of matrices to which Corollary 1 applies. Thus stability of the family G_N implies that $A(x)$ has no zero inside the unit circle unless it is also a zero of $A(1/x)$. But a common zero of $A(x)$ and $A(1/x)$ would imply that D is singular, which would contradict property (4). Therefore $A(x)$ has no zero inside the unit circle. Stability implies further that (11.12) holds, where M is defined as before, and this implies in turn that $M \leq 2$, which, in view of (11.11), is tantamount to (11.6) and therefore to (11.3). ∎

It is easily verified that G^{∞}, when it exists, is the orthogonal projector on the eigenspace of G associated with the eigenvalue 1, that is the space of N-vectors whose components are successive equally spaced ordinates of polynomials of degree $s - 1$ or less.

There is a curious unsolved mathematical problem concerning the stability theorem. It will be recalled that the symmetry of G was not included in the hypothesis of Theorem 5.1. Rather this was a consequence of the general hypotheses and the five defining properties. However, in Theorem 11.1 the symmetry of G is hypothesized. While symmetry of the main part of G follows from the symmetry of the coefficients in the main formula and properties (1) to (4), the special corner submatrices are not symmetric unless $A(x)$ is chosen so that $B(x) = A(x)$. When the characteristic function of the given

MWA satisfies (11.3), we might wish to replace property (5) by the requirement that the family G_N be stable, and still hope to have G uniquely determined. At present this appears to require the additional hypothesis that G be symmetric, because the proof of stability (Greville 1980) involves extensive use of the well known relation between Rayleigh quotients and eigenvalues that holds only for Hermitian (including symmetric) matrices. If $q(z)$ has a number of zeros (none, we assume, on the unit circle), there are, in general, some possible extensions with unsymmetric corners. I conjecture that there is, in such a case, no unsymmetric stable extension, but I have not been able to prove this; nor have I been able to find a counter-example to the conjecture. Thus the possibility exists (though I think it unlikely) that some symmetrical MWA (with $q_o > 0$ and no zeros on the unit circle) might have more than one stable extension, the unique symmetric one and an unsymmetric one as well.

It may be mentioned, however, that there are cogent reasons for thinking that G should be symmetric. A square matrix is called persymmetric if it is symmetric about its secondary diagonal. It is called centrosymmetric if it is symmetric about the center of the matrix: thus $C = (c_{ij})$ is centrosymmetric if $c_{ij} = c_{N-j+1,N-i+1}$ for all (i,j). Now, it is easily seen that of the three properties of symmetry, persymmetry, and centrosymmetry, any two imply the third. G is necessarily persymmetric, because $G = I - F$, where F is a Trench matrix, and every Trench matrix is persymmetric (see Greville and Trench 1979). Therefore, if G is not symmetric, it is not centrosymmetric. Now, if G is not centrosymmetric, this means that reversing the order of the observed values would not merely reverse the order of the smoothed values, but would cause different numerical values to be obtained. For example, the elements of the bottom row of G would not be those of the top row in reverse

order. The formula for smoothing the last observation would not be the mirror image of the one for smoothing the first observation, but would be a different formula. This would seem to be an undesirable characteristic of the smoothing process.

12. SMOOTHING FORMULAS IN THE STRICT SENSE AND AN OPTIMAL PROPERTY

Under certain conditions the smoothing procedure described herein can be shown to minimize a certain "loss function" analogous to the Whittaker criterion. In a slightly more general form of the Whittaker smoothing method (Greville 1957) one minimizes the sum of the squares of the departures of the smoothed values from the observed values plus a specified quadratic form in the sth differences of the smoothed values. In matrix terms, one minimizes

$$(u - y)^T (u - y) + (Ku)^T HKu \ , \qquad (12.1)$$

where H is a given positive definite matrix of order $N - s$. Minimization of this expression leads to the equation

$$(I + K^T HK)u = y \ ,$$

which has a unique solution for u since $I + K^T HK$ is positive definite. I showed (Greville 1957) that this graduation method has the interesting property that if roughness (opposite of smoothness) is measured by the term $(Ku)^T HKu$, smoothness is always increased by the graduation. By Theorem 5.22 of Noble (1969),

$$(I + K^T HK)^{-1} = I - K^T(H^{-1} + KK^T)^{-1} K.$$

The last expression is of the form (5.2) and suggests that the use of an MWA with the natural extension might be regarded as a generalized Whittaker smoothing process if

$$D = (H^{-1} + KK^T)^{-1} \ .$$

Solving for H gives

$$H = (D^{-1} - KK^T)^{-1} \ . \qquad (12.2)$$

We are led to inquire, therefore, under what conditions an MWA is such that the right member of (12.2) for the natural extension is positive definite for all N . We note in passing that

$$H^{-1} = D^{-1} - KK^T$$

is a Toeplitz matrix.

Schoenberg (1946, p. 53) remarks that it is desirable for an efficient smoothing formula, one that achieves adequate smoothness without producing unnecessarily large departures from the observed values, to have its characteristic function satisfy the stronger condition

$$0 \leq \phi(t) \leq 1 \tag{12.3}$$

for all t . This remark seems to have been little noted in the years since its publication. We shall call an MWA a <u>smoothing formula in the strict sense</u> if its characteristic function satisfies (12.3).

<u>Lemma 12.1.</u> Under the natural extension of a given MWA, $D^{-1} - KK^T$ is nonsingular if and only if G is nonsingular, and H defined by (12.2) is positive definite if and only if G is positive definite.

<u>Proof</u>. If

$$G = I - K^T DK , \tag{12.4}$$

as in (5.2), then by Noble's theorem

$$G^{-1} = I + K^T(D^{-1} - KK^T)^{-1} K , \tag{12.5}$$

provided G and D are nonsingular. Under the natural extension, D is always nonsingular by property (4). In the proof of Noble's theorem, the nonsingularity of $D^{-1} - KK^T$ is shown to follow from that of G and D . On the other hand, if $D^{-1} - KK^T$ is nonsingular, multiplication of the right members of (12.4) and (12.5) gives the identity. This proves the first statement of the lemma.

Now let H be positive definite. We have shown that

$$G^{-1} = I + K^T HK .$$

Then, if v is an arbitrary nonzero real vector,

$$v^T G^{-1} v = v^T v + (Kv)^T HKv. \tag{12.6}$$

The second term of the right member of (12.6) is nonnegative, since H is positive definite, and the first term is positive. It follows that G^{-1}, and therefore G , is positive definite.

Conversely, let G be positive definite. Applying Noble's theorem to (12.2) gives

$$H = D + DK(I - K^T DK)^{-1} K^T D = D + DKG^{-1} K^T D.$$

Now, we note that under the natural extension D is positive definite (Greville 1980, Theorem 1), since all the zeros of A(x) are outside the unit circle. Thus, the same argument used previously shows that $v^T Hv > 0$ for every nonzero real vector v , and so H is positive definite. ∎

__Theorem 12.2.__ Under the natural extension of a given MWA, H is positive definite for all N if and only if $\phi(t)$ satisfies (12.3).

__Proof.__ By Lemma 12.1, H is positive definite if and only if G is positive definite; therefore we need consider only the positive definiteness of G . We recall that G = I - F , where F is a singular, symmetric Trench matrix characterized by two identical polynomials equal to $\hat{A}(x)$. Since all the zeros of $\hat{A}(x)$, with the exception of +1 , are outside the unit circle, F is positive semidefinite (Greville 1980, Theorem 1), and if

$$f(x) = \hat{A}(x)\ \hat{A}(1/x),$$

then

$$\psi(t) = f(e^{it}) = |A(e^{it})|^2$$

is nonnegative for all t . Let M denote the maximum of $\psi(t)$.

Let $\phi(t)$ satisfy (12.3). Since $\phi(t) = 1 - \psi(t)$, it follows that

$$0 \leq \psi(t) \leq 1 \qquad (12.7)$$

for all t. Therefore $M \leq 1$, and it follows (Greville 1980, Theorem 2) that for all N all eigenvalues of F are nonnegative and less than unity. Since the eigenvalues of G are 1 minus those of F, all of the former are positive for all N, and therefore G is positive definite for all N.

Conversely, let G be positive definite for all N. Then all its eigenvalues are positive for all N, and consequently those of F are less than unity (but not less than zero, since F is positive semidefinite). Since M is the limit of the largest eigenvalue as N approaches infinity (Greville 1980, Theorem 2), $M \leq 1$. Therefore (12.7) holds, and it is equivalent to (12.3). ∎

It is easy to construct an MWA that is a smoothing formula in the strict sense. However, none of the weighted averages in common use fall in this class. As a practical matter, the smoothing effected by such formulas is likely to be too "gentle." In particular, using the properties of Jacobi polynomials, I have shown elsewhere (Greville 1966) that the characteristic functions of all minimum-R_s averages assume negative values in $(0, 2\pi)$. Thus no such formula is a smoothing formula in the strict sense.

There is, however, one family of moving averages, mentioned in the literature but not in general use, that are smoothing formulas in the strict sense. This is the limiting case of the minimum-R_s formulas as s approaches infinity (Greville 1966). In finite-difference form, the minimum-R_∞ MWA of $2m + 1$ terms, exact for the degree $2s - 1$, is

$$u_x = \mu^{2(m-s+1)} \sum_{j=0}^{s-1} (-4)^{-j} \binom{m-s+j}{j} \delta^{2j} y_x ,$$

where the operator μ is defined by

$$\mu f(x) = \tfrac{1}{2}[f(x + \tfrac{1}{2}) + f(x - \tfrac{1}{2})] ,$$

so that $\mu^2 = 1 + \tfrac{1}{4}\delta^2$. The characteristic function is

$$\phi(t) = (\cos \tfrac{1}{2}t)^{2(m-s+1)} \sum_{j=0}^{s-1} \binom{m-s+j}{j} \sin^{2j} \tfrac{1}{2}t,$$

which is nonnegative in $0 < t < 2\pi$, with a single zero of multiplicity $2(m - s + 1)$ at $t = \pi$.

It may be mentioned that, in the case where $\phi(t)$ assumes some negative values (and G and H are nonsingular), though the expression (12.1) does not have an extremum, the natural extension of the graduation does correspond to a saddle point of (12.1). It is not clear what significance this observation may have.

ACKNOWLEDGMENTS

This paper has benefited from discussions with many persons, but I wish to thank especially D. R. Schuette for invaluable help with the computations, and J. M. Hoem and W. F. Trench for their careful reading of the manuscript, which led to significant improvements. I am solely responsible for any errors that may be found.

1. Monthly Precipitation (Inches), Madison, Wisconsin, 1967-71.

Year and Month		Observed Value	Graduated Value	Year and Month		Observed Value	Graduated Value
1967	January	1.63	1.11	1969	July	4.28	3.81
	February	1.17	1.63		August	0.96	3.17
	March	1.49	2.24		September	1.35	2.33
	April	2.57	2.88		October	2.65	1.56
	May	3.53	3.42		November	0.70	1.06
	June	6.46	3.74		December	1.66	0.82
	July	2.51	3.85	1970	January	0.44	0.90
	August	2.71	3.75		February	0.16	1.25
	September	2.68	3.42		March	1.17	1.78
	October	5.52	2.92		April	2.53	2.39
	November	1.83	2.31		May	6.09	2.94
	December	1.89	1.69		June	2.26	3.37
1968	January	0.56	1.31		July	2.42	3.63
	February	0.49	1.36		August	0.97	3.69
	March	0.59	1.87		September	8.82	3.50
	April	4.18	2.69		October	2.65	3.20
	May	2.02	3.49		November	1.06	2.74
	June	7.82	3.91		December	2.12	2.28
	July	2.54	3.92	1971	January	1.48	1.94
	August	2.58	3.54		February	2.59	1.76
	September	4.45	2.97		March	1.52	1.74
	October	0.85	2.45		April	2.42	1.81
	November	1.74	1.99		May	0.98	1.93
	December	2.89	1.64		June	2.27	2.02
1969	January	2.26	1.56		July	1.65	2.13
	February	0.18	1.81		August	3.96	2.24
	March	1.47	2.35		September	1.87	2.40
	April	2.72	3.13		October	1.30	2.63
	May	3.45	3.81		November	3.48	2.84
	June	7.96	4.05		December	3.64	3.28

SOURCE: Observed values from U. S. Department of Commerce, National Oceanic and Atmospheric Administration, Environmental Data Service, Local Climatological Data. Annual Summary with Comparative Data. Madison, Wisconsin, 1972, National Climatic Center, Asheville, N. C., 1973.

2. Moving-Average Coefficients[a] (c_j) and Extension Coefficients (a_j) of Minimum-R_3 ("Henderson's Ideal") Averages of 5 to 23 Terms Exact for Cubics

	Number of Terms									
	5		7		9		11		13	
j	c_j	a_j	c_j	a_j	c_j	a_j	c_j	a_j	c_j	a_j
0	.559440		.412588		.331140		.277944		.240058	
1	.293706	2	.293706	1.618034	.266557	1.352613	.238693	1.160811	.214337	1.016301
2	-.073426	-1	.058741	-.236068	.118470	.114696	.141268	.281079	.147356	.360880
3			-.058741	-.381966	-.009873	-.287231	.035723	-.140968	.065492	-.021625
4					-.040724	-.180078	-.026792	-.204545	0	-.160909
5							-.027864	-.096377	-.027864	-.138330
6									-.019350	-.056317

[a]Calculated by formula (9.5).

2. Moving-Average Coefficients[a] (c_j) and Extension Coefficients (a_j) of Minimum-R_3 ("Henderson's Ideal") Averages of 5 to 23 Terms Exact for Cubics (continued)

	Number of Terms									
	15		17		19		21		23	
j	c_j	a_j	c_j	a_j	c_j	a_j	c_j	a_j	c_j	a_j
0	.211542		.189232		.171266		.156470		.144060	
1	.193742	.903661	.176390	.813444	.161691	.739580	.149136	.678000	.138318	.625880
2	.145904	.397295	.141112	.410885	.134965	.412090	.128423	.406495	.121949	.397207
3	.082918	.064751	.092293	.124932	.096658	.166162	.097956	.193174	.097395	.212501
4	.024028	-.100710	.042093	-.043456	.054685	.005097	.063038	.046016	.068303	.075236
5	-.014134	-.135445	.002467	-.110644	.017474	-.078255	.029628	-.046290	.038933	-.015313
6	-.024499	-.094424	-.018640	-.106213	-.008155	-.099972	.003119	-.084020	.013430	-.063927
7	-.013730	-.035128	-.020370	-.065896	-.018972	-.081843	-.012896	-.084711	-.004948	-.078737
8			-.009961	-.023052	-.016601	-.047103	-.017614	-.063086	-.014527	-.070064
9					-.007378	-.015756	-.013455	-.034444	-.015687	-.048977
10							-.005570	-.011134	-.010918	-.025714
11									-.004278	-.008092

[a]Calculated by formula (9.5).

3. **Moving-Average Coefficients (c_j) and Extension Coefficients (a_j) of Selected Moving Averages**

	Macaulay[a]		Spencer 15-Term[b]		Woolhouse[c]		Hardy[d]		Higham[e]		Karup[f]	
j	$864c_j$	a_j	$320c_j$	a_j	$125c_j$	a_j	$120c_j$	a_j	$125c_j$	a_j	$625c_j$	a_j
0	182		74		25		24		25		125	
1	171	.919760	67	.961572	24	.885108	22	.739988	24	.859550	114	.820240
2	127	.393023	46	.372752	21	.421982	17	.386211	18	.399283	87	.402924
3	72	.055273	21	.015904	7	.028721	10	.124325	10	.087040	53	.114622
4	17	-.113111	3	-.123488	3	-.076050	4	-.023648	3	-.072738	21	-.047133
5	-17	-.140462	-5	-.125229	0	-.107285	0	-.080087	0	-.104527	0	-.102491
6	-19	-.084512	-6	-.075887	-2	-.092723	-2	-.079459	-2	-.093953	-8	-.091791
7	-10	-.029971	-3	-.025624	-3	-.059753	-2	-.049327	-2	-.055312	-9	-.060239
8							-1	-.018003	-1	-.019343	-6	-.028636
9											-2	-.007496

[a] Macaulay 1931, p. 55, footnote 2.

[b] Macaulay 1931, p. 55; Henderson 1938, p. 53.

[c] Henderson 1938, p. 53.

[d] Henderson 1938, p. 53; Benjamin and Haycocks 1970, p. 238.

[e] Henderson 1938, p. 53.

[f] Henderson 1938, p. 53.

3. Moving-Average Coefficients (c_j) and Extension Coefficients (a_j) of Selected Moving Averages (continued)

	Andrews[g]		Spencer 21-Term[h]		Hardy Wave-Cutting[i]		Vaughan Formula A[j]		Kenchington[k]	
j	$10080c_j$	a_j	$350c_j$	a_j	$65c_j$	a_j	$1440c_j$	a_j	$385c_j$	a_j
0	1688		60		5		182		45	
1	1579	.700747	57	.729724	5	.480996	179	.593256	44	.527740
2	1325	.406808	47	.408707	6	.368708	170	.396409	41	.370688
3	950	.179749	33	.167281	7	.267940	149	.230238	36	.236445
4	551	.027155	18	.009255	7	.166506	115	.096761	30	.128638
5	225	-.054586	6	-.069703	6	.072964	72	-.000857	22	.043118
6	-4	-.083701	-2	-.091513	4	-.008222	29	-.060076	13	-.018390
7	-124	-.078256	-5	-.076165	1	-.075454	-5	-.083321	5	-.053902
8	-135	-.054368	-5	-.049051	-1	-.097387	-26	-.079596	-1	-.067080
9	-110	-.031120	-3	-.022502	-2	-.089039	-29	-.056662	-5	-.064844
10	-61	-.012428	-1	-.006033	-2	-.062016	-19	-.028557	-6	-.050323
11					-1	-.024996	-6	-.007595	-5	-.032035
12									-3	-.015626
13									-1	-.004429

[g]Andrews and Nesbitt 1965, p. 18.

[h]Macaulay 1931, p. 51; Henderson 1938, p. 53.

[i]Benjamin and Haycocks 1970, p. 239.

[j]Vaughan 1933, p. 437.

[k]Henderson 1938, p. 53.

4. Matrix Elements for the Natural Extension of the 9-Term Minimum-R_3 Smoothing Formula, with Second-Difference Correction and R_o Value for Each Supplemental Formula

i	j = 1	2	3	4	5	6	7	8	9	Second-Difference Correction	R_o
1	.773854	.305888	.025938	-.064956	-.040724	0	0	0	0	-.4133	.8360
2	.305888	.360106	.270804	.113799	-.009873	-.040724	0	0	0	.1457	.5579
3	.025938	.270804	.357131	.278254	.118470	-.009873	-.040724	0	0	.1931	.5429
4	-.064956	.113799	.278254	.338473	.266557	.118470	-.009873	-.040724	0	.0744	.5441
5	-.040724	-.009873	.118470	.266557	.331140	.266557	.118470	-.009873	-.040724	0	.5322

5. Matrix Elements for the Least-Squares Extension of the 9-Term Minimum-R_3 Smoothing Formula with Fourth Difference Correction and R_o Value for Each Supplemental Formula

i	j = 1	2	3	4	5	6	7	8	9	Fourth-Difference Correction	R_o
1	.335350	.058600	-.037900	.058600	-.014650	0	0	0	0	-.01465	.9928
2	.025386	.857731	.315214	-.345889	.188282	-.040724	0	0	0	-.01534	.9962
3	-.206335	.652571	.412341	.048375	.240395	-.009873	-.040724	0	0	-.41580	.8369
4	-.140189	.232497	.299136	.241547	.299136	.118470	-.009873	-.040724	0	-.68194	.5717
5	-.040724	-.009873	.118470	.266557	.331140	.266557	.118470	-.009873	-.040724	-.75525	.5322

6. **Third Differences of Matrix Elements for the Natural and Least-Squares Extensions of the 9-Term Minimum-R_3 Smoothing Formula**

i	j = 1	2	3	4	5	6	7	8	9
					Natural Extension				
1	-.000046	.075817	-.006665	-.064956	-.077748	.025917	.112299	-.040724	0
2	.130190	.101036	.084297	-.027899	-.103248	-.077748	.025917	.112299	-.040724
3	-.098634	.059488	.112348	.055964	-.045662	-.103248	-.077748	.025917	.112299
4	-.057216	-.021246	.066051	.095915	.045662	-.045662	-.103248	-.077748	.025917
5	.040724	-.112299	-.025917	.077748	.103248	.045662	-.045662	-.103248	-.077748
					Least-Squares Extension				
1	-.430376	.789377	.095655	-.709595	.157447	.025917	.112299	-.040724	0
2	-.264548	.392618	.142871	-.257320	-.033365	-.077748	.025917	.112299	-.040724
3	-.092060	.033815	.119784	.091815	-.069850	-.103248	-.077748	.025917	.112299
4	.018017	-.139944	.045169	.192841	.013083	-.045662	-.103248	-.077748	.025917
5	.040724	-.112299	-.025917	.077748	.103248	.045662	-.045662	-.103248	-.077748

7. Extension of 15-Term Spencer Graduation of Madison Precipitation Data to First Seven and Last Seven Months by Different Methods

Year and Month	Observed Value	Extension of Graduation by		
		Natural Method	Least-Squares Cubic	Additional Data
1967				
January	1.63	1.11	1.62	1.56
February	1.17	1.63	0.98	1.84
March	1.49	2.24	1.37	2.29
April	2.57	2.88	2.32	2.85
May	3.53	3.42	3.07	3.36
June	6.46	3.74	3.61	3.70
July	2.51	3.85	3.82	3.84
- - -	- - -	- - -	- - -	- - -
1971				
June	2.27	2.02	2.00	2.05
July	1.65	2.13	2.03	2.23
August	3.96	2.24	2.00	2.39
September	1.87	2.40	1.97	2.51
October	1.30	2.63	2.08	2.50
November	3.48	2.84	2.58	2.31
December	3.64	3.28	3.85	2.04

8. Parameters of the Symmetrical Weighted Averages Listed in Tables 2 and 3

Designation	Number of Terms	R_o	R_3	Error
Minimum-R_3 (Henderson's ideal):	5	.7045	.2735	$-.073\delta^4$
	7	.5971	.1147	$-.29\delta^4$
	9	.5323	.0581	$-.76\delta^4$
	11	.4865	.0331	$-1.57\delta^4$
	13	.4515	.0204	$-2.88\delta^4$
	15	.4234	.0134	$-4.85\delta^4$
	17	.4002	.0095	$-7.64\delta^4$
	19	.3806	.0066	$-11.4\delta^4$
	21	.3636	.0048	$-16.5\delta^4$
	23	.3488	.0036	$-23.1\delta^4$
Macaulay	15	.4273	.01657	$-4.52\delta^4$
Spencer	15	.4389	.01659	$-3.86\delta^4$
Woolhouse	15	.4602	.0654	$-5.4\delta^4$
Hardy	17	.4059	.0105	$\frac{1}{12}\delta^2 - 3.70\delta^4$
Higham	17	.4127	.0179	$-6.4\delta^4$
Karup	19	.4036	.0095	$-7.8\delta^4$
Andrews	21	.3707	.00628	$-14.9\delta^4$
Spencer	21	.3784	.00626	$-12.6\delta^4$
Hardy, wave-cutting	23	.3332	.0154	$-48.8\delta^4$
Vaughan A	23	.3415	.0050	$-26.6\delta^4$
Kenchington	27	.3202	.0031	$-22.4\delta^4$

Figure A. Observed and Graduated Values of Monthly Precipitation, Madison, Wisconsin, 1967-1971

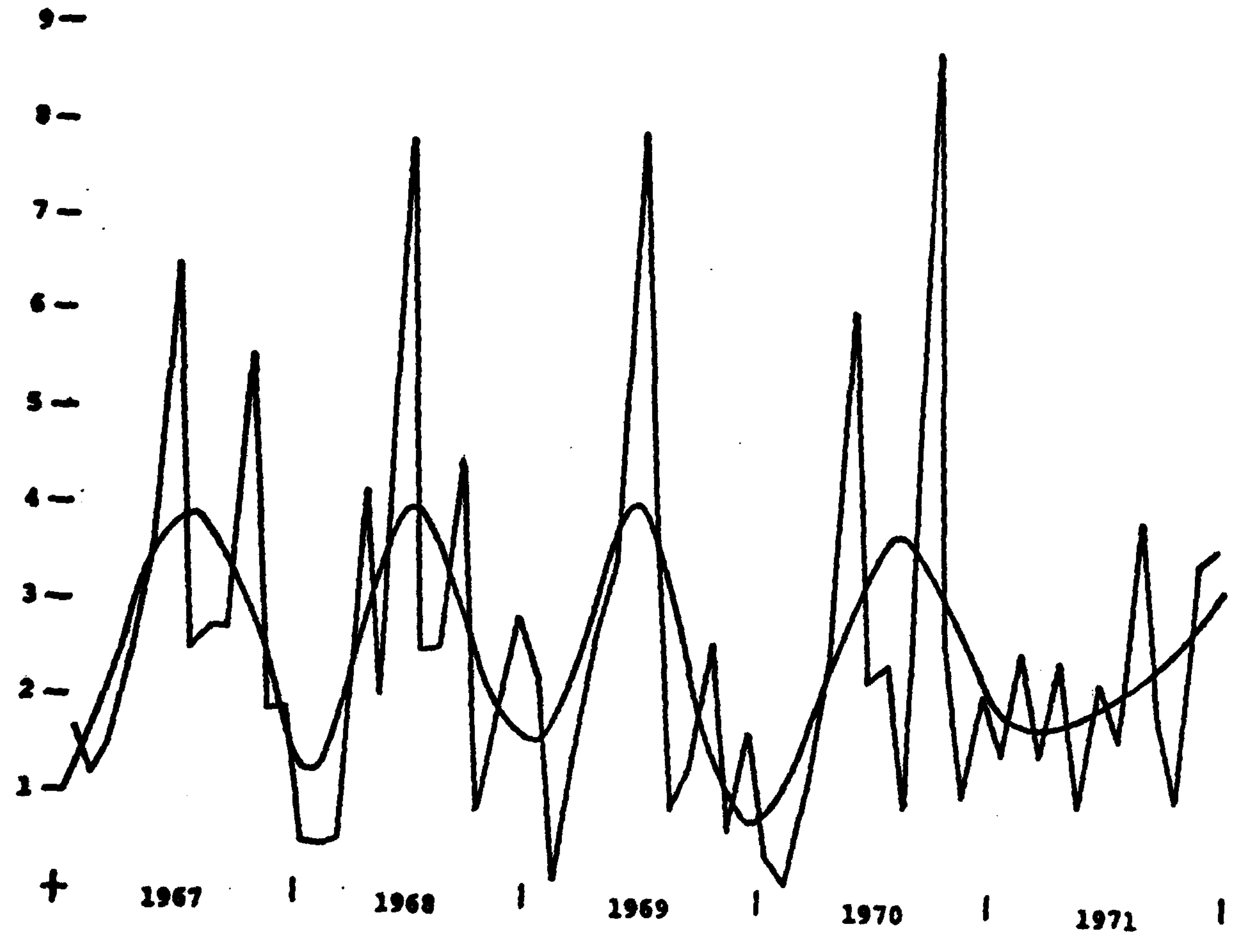


Figure 8. Observed and Graduated Values of Monthly Precipitation, Madison, Wisconsin, January to March, 1967

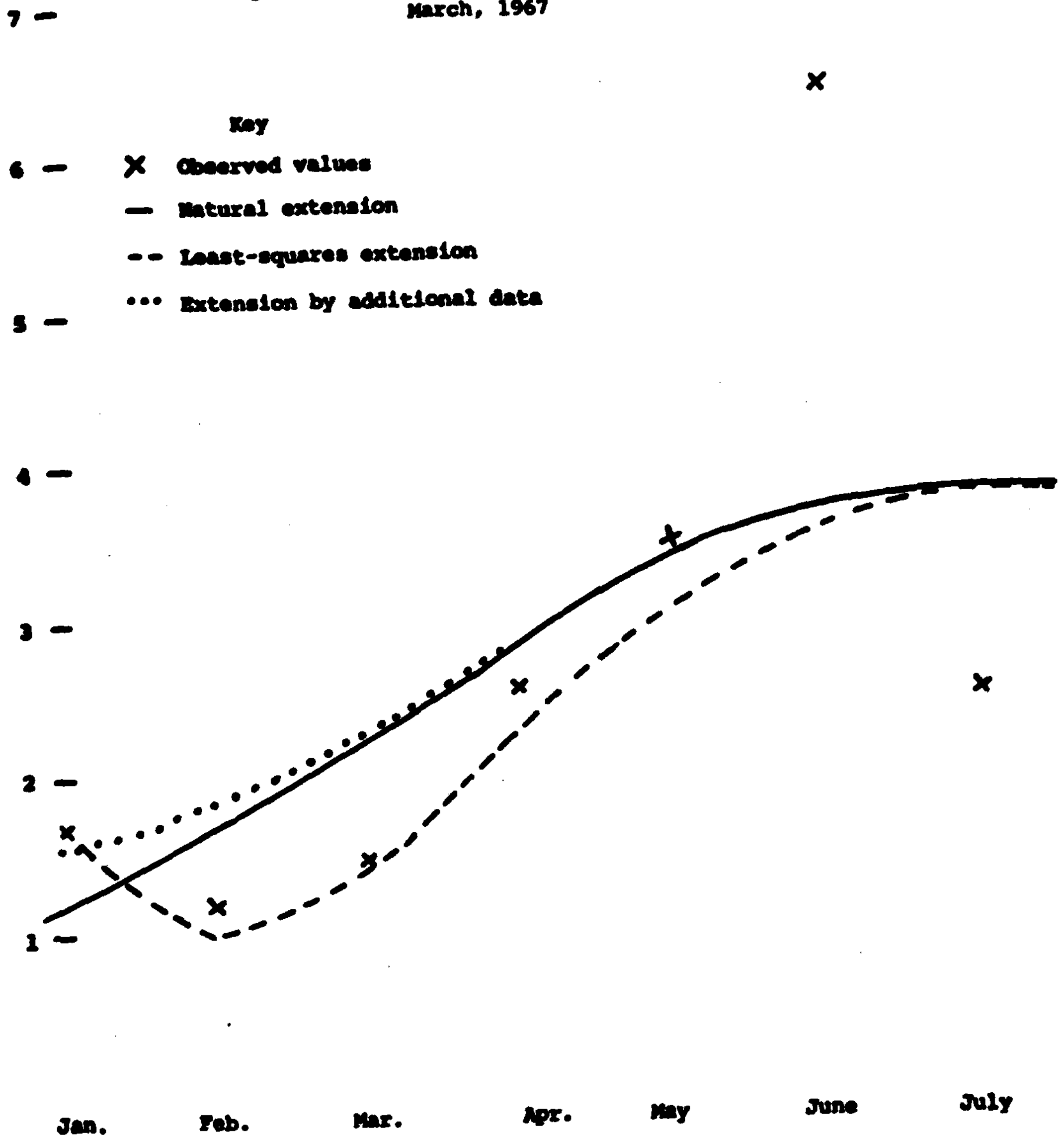
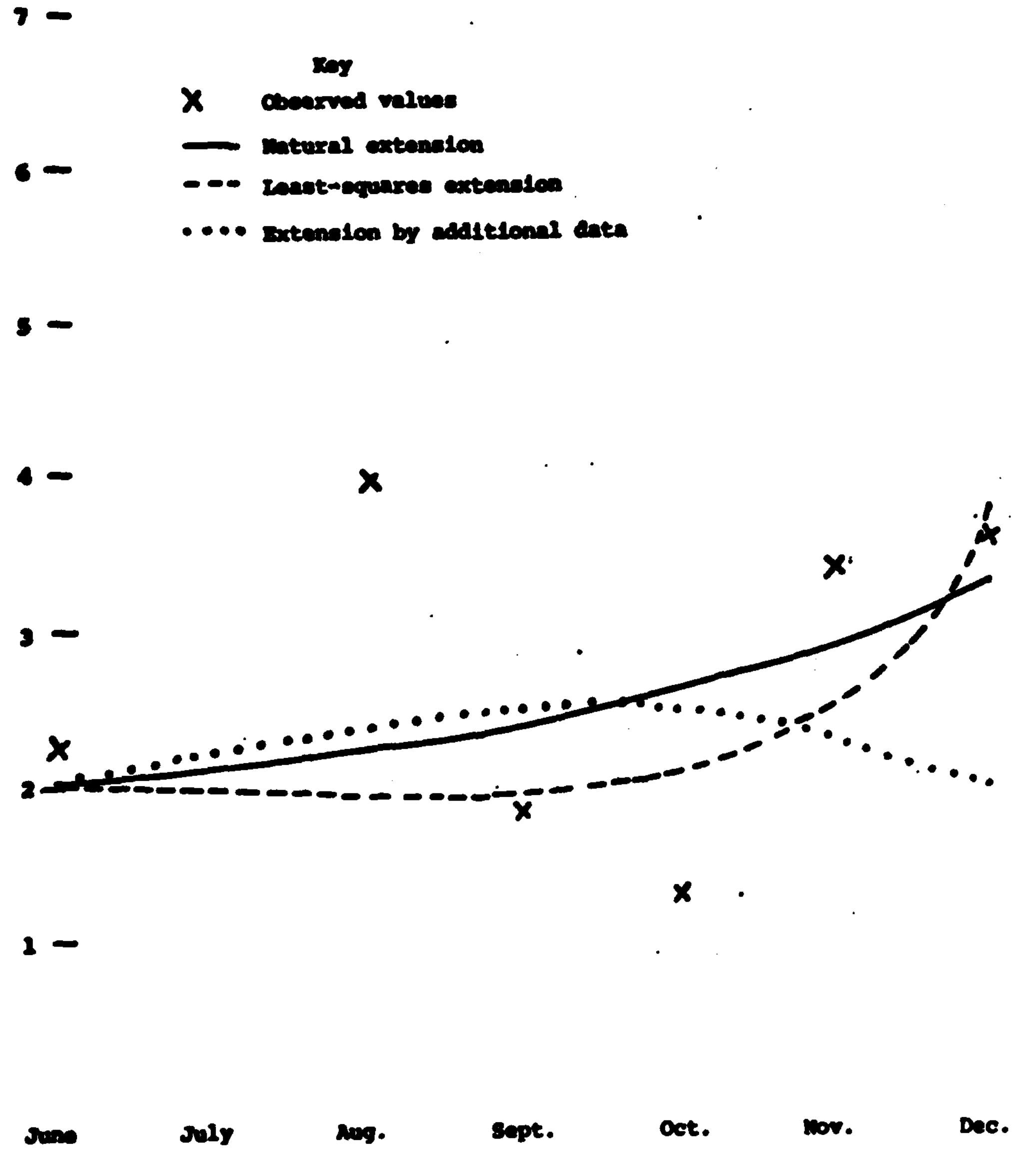


Figure C. Observed and Graduated Values of Monthly Precipitation, Madison, Wisconsin, July to December, 1971



REFERENCES

Aitken, A. C. (1926), "The Accurate Solution of the Difference Equation Involved in Whittaker's Method of Graduation, and its Practical Application," Transactions of the Faculty of Actuaries, 11, 31-39.

Andrews, George H., and Nesbitt, Cecil J. (1965), "Periodograms of Graduation Operators," Transactions of the Society of Actuaries, 17, 1-27.

Ben-Israel, Adi, and Greville, Thomas N. E. (1974), Generalized Inverses: Theory and Applications, New York: John Wiley and Sons.

Benjamin, Bernard, and Haycocks, H. W. (1970), The Analysis of Mortality and Other Actuarial Statistics, Cambridge, Eng.: Cambridge University Press.

De Forest, Erastus L. (1873), "On Some Methods of Interpolation Applicable to the Graduation of Irregular Series, Such as Tables of Mortality, &c., &c.," Smithsonian Report 1871, 275-339.

—— (1875), "Additions to a Memoir on Methods of Interpolation Applicable to the Graduation of Irregular Series," Smithsonian Report 1873, 319-349.

—— (1876), Additions to a Memoir on Methods of Interpolation Applicable to the Graduation or Adjustment of Irregular Series of Observed Numbers, New Haven: Tuttle, Morehouse, and Taylor Co.

—— (1877), "On Adjustment Formulas," The Analyst, 4, 79-86 and 107-113.

Dresden, Arnold (1942), "On the Iteration of Linear Homogeneous Transformations," Bulletin of the American Mathematical Society, 48, 577-579.

Elphinstone, M. D. W. (1951), "Summation and Some Other Methods of Graduation —the Foundations of Theory," Transactions of the Faculty of Actuaries, 20, 15-77.

Grenander, Ulf, and Szegö, Gabor (1958), Toeplitz Forms and Their Applications, Berkeley: Univ. of Calif. Press.

Greville, Thomas N. E. (1947), "Actuarial Note: Adjusted Average Graduation Formulas of Maximum Smoothness," Record of the American Institute of Actuaries, 36, 249-264.

—— (1948), "Actuarial Note: Tables of Coefficients in Adjusted Average Graduation Formulas of Maximum Smoothness," Record of the American Institute of Actuaries, 37, 11-30.

—— (1957), "On Smoothing a Finite Table: a Matrix Approach," Journal of the Society for Industrial and Applied Mathematics, 5, 137-154.

—— (1966), "On Stability of Linear Smoothing Formulas," SIAM Journal on Numerical Analysis, 3, 157-170.

—— (1974a), Part 5 Study Notes, Graduation (1974 edition), Chicago: Education and Examination Committee of the Society of Actuaries.

—— (1974b), "On a Problem of E. L. De Forest in Iterated Smoothing," SIAM Journal on Mathematical Analysis, 5, 376-398.

—— (1979), "On a Problem Concerning Band Matrices with Toeplitz Inverses," Proceedings of the Eighth Manitoba Conference on Numerical Mathematics and Computing, 275-283, Winnipeg: Utilitas Mathematica Publishing Inc.

—— (1980), "Bounds for Eigenvalues of Hermitian Trench Matrices," to appear in Proceedings of the Ninth Manitoba Conference on Numerical Mathematics and Computing, Winnipeg: Utilitas Mathematica Publishing Inc.

—— and Trench, William F. (1979), "Band Matrices with Toeplitz Inverses," Linear Algebra and its Applications, 27, 199-209.

Hannan, E. J. (1970), Multiple Time Series, New York: John Wiley and Sons.

Hardy, George F. (1909), The Theory of Conaruction of Tables of Mortality and of Similar Statistical Tables in Use by the Actuary, London: Institute of Actuaries.

Henderson, Robert (1916), "Graduation by Adjusted Average," Transactions of the Actuarial Society of America, 17, 43-48.

——— (1924), "A New Method of Graduation," Transactions of the Actuarial Society of America, 25, 29-40.

——— (1938), Mathematical Theory of Graduation, New York: Actuarial Society of America.

Macaulay, Frederick R. (1931), The Smoothing of Time Series, New York: National Bureau of Economic Research.

Maclean, Joseph B. (1913), "Graduation by the Summation Method, Some Elementary Notes," Transactions of the Actuarial Society of America, 14, 256-276.

Miller, Morton D. (1946), Elements of Graduation, New York and Chicago: Actuarial Society of America and American Institute of Actuaries.

Noble, Ben (1969), Applied Linear Algebra, Englewood Cliffs, N. J.: Prentice-Hall, Inc.

Oldenburger, Rufus (1940), "Infinite Powers of Matrices and Characteristic Roots," Duke Mathematical Journal, 6, 357-361.

Schiaparelli, G. V. (1866), "Sul Modo di Ricavare la Vera Espressione delle Legge della Natura dalle Curve Empiriche," appendix to Effemeredi Astronomiche di Milano per l'Anno 1867, Milan.

Schoenberg, I. J. (1946), "Contributions to the Problem of Approximation of Equidistant Data by Analytic Functions," Quarterly of Applied Mathematics, 4, 45-99 and 112-141.

——— (1948), "Some Analytical Aspects of the Problem of Smoothing," in Studies and Essays Presented to R. Courant on His 60th Birthday, New York: Interscience Publishers, 351-370.

——— (1953), "On Smoothing Operations and Their Generating Functions," Bulletin of the American Mathematical Society, 59, 199-230.

Sheppard, W. F. (1913), "Reduction of Errors by Means of Negligible Differences," Proceedings of the Fifth International Congress of Mathematicians, 2, 348-384.

Shiskin, Julius, and Eisenpress, Harry (1957), "Seasonal Adjustments by Electronic Computer Methods," Journal of the American Statistical Association, 52, 415-449.

——, Young, Allan H., and Musgrave, John C. (1967), The X-11 Variant of the Census Method II Seasonal Adjustment Program, Washington: U. S. Department of Commerce, Bureau of the Census.

Thomée, Vidar (1965), "Stability of Difference Schemes in the Maximum-Norm," Journal of Differential Equations, 1, 273-292.

Thrall, Robert M., and Tornheim, Leonard (1957), Vector Spaces and Matrices, New York: John Wiley and Sons.

Vaughan, Hubert (1933), "Summation Formulas of Graduation with a Special Type of Operator," Journal of the Institute of Actuaries, 64, 428-448.

Whittaker, Edmund T. (1923), "On a New Method of Graduation," Proceedings of the Edinburgh Mathematical Society, 41, 63-75.

Wiener, Norbert (1949), Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications, New York: John Wiley and Sons.

Wolfenden, Hugh H. (1925), "on the Development of Formulae for Graduation by Linear Compounding, with Special Reference to the Work of Erastus L. De Forest," Transactions of the Actuarial Society of America, 26, 81-121.

—— (1942), The Fundamental Principles of Mathematical Statistics, Toronto: Macmillan Co. of Canada, Ltd.

TNEG/db

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 2025	2. GOVT ACCESSION NO. AD-A083 808	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) MOVING-WEIGHTED-AVERAGE SMOOTHING EXTENDED TO THE EXTREMITIES OF THE DATA		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) T. N. E. Greville		8. CONTRACT OR GRANT NUMBER(s) DAAG29-75-C-0024
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 3 - Numerical Analysis and Computer Science
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE November 1979
		13. NUMBER OF PAGES 64
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Smoothing; Moving weighted averages; Extrapolation; Toeplitz matrices; Graduation.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The use of a symmetrical moving weighted average of $2m + 1$ terms to smooth equally spaced observations of a function of one variable does not yield smoothed values of the first m and the last m observations, unless additional data beyond the range of the original observations are available. By means of analogies to the Whittaker smoothing process and some related mathematical concepts, a natural method is developed for extending the smoothing to the extremities of the data as a single overall matrix-vector operation having a well defined structure, rather than as something extra grafted on at the ends.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)